A robust structural electric system model with significant share of intermittent renewables under auto-correlated residual demand

Pierre Cayet
Arash Farnoosh

2022-6 Document de Travail/ Working Paper

EconomiX - UMR 7235 Bâtiment Maurice Allais
Université Paris Nanterre 200, Avenue de la République
92001 Nanterre Cedex

Site Web : economix.fr
Contact : secreteriat@economix.fr
Twitter : @EconomixU

A robust structural electric system model with significant share of intermittent renewables under auto-correlated residual demand

Pierre CAYET[a]*, Arash FARNOOSH[a]

* Corresponding author: *Email address*: pierre.cayet@ifpen.fr ; *postal address:* IFP School, 232 Avenue Napoléon Bonaparte, 92852 Rueil-Malmaison, France

a. IFPEN, IFP School, 232 Avenue Napoléon Bonaparte, 92852 Rueil-Malmaison, France

***Abstract* – In this paper, we propose a robust structural investment and dispatch model of electric systems including commitment and storage constraints under auto-correlated residual demand. We associate it to a novel approach to robust optimization focusing on uncertainty parameter trajectories. Using Principal Component Analysis, we approximate conditional order statistics for the differential distribution of components of residual demand using parametric polynomial approximation. This flexible method allows us to derive a set of extreme trajectories maximizing the level and variability of residual demand. Finally, we apply our dynamic robust model to the electric system of the French region Auvergne Rhône-Alpes and discuss the implications in terms of investment decisions and cost performance.**

***Index Terms* –** Optimal electricity mix; Robust optimization; Dynamic uncertainty; Renewable energy

I. INTRODUCTION

Robust optimization provides a natural modeling technique for quantifying uncertainties affecting future weather and electric systems. Increasing penetration of renewables generates higher volatility of residual demand [1], which in turns requires higher system flexibility through technical solutions including battery and electric vehicle storage, gas fired power plants, demand-side management (DSM) and curtailment [2]. As underlined by [3], reserves are increasingly used to cover fluctuations of power output as the share of renewables increase, which requires stochastic decision-making tools. However, in the absence of reliable probabilistic description of the joint distribution of residual demand at various point of time and locations in the grid, robust optimization provides a non-probabilistic tool allowing to minimize the dispatch and recourse cost under the worst-case realization of the uncertainty.

First introduced by [4], robust-optimization provides a non-probabilistic formulation of uncertainty [5]. The uncertainty set, defined in the real space, is the set of values the uncertain parameters can take [6]-[7]. Recent developments introduce correlation between uncertain parameters [8]-[9] and dynamic uncertainty sets for multi-period optimization problems [10]. For each time period, the value of the uncertain parameter determines the shape and size of the uncertainty set corresponding to the following period. Similarly to [11], our method allows us to model distributional asymmetries of the uncertain parameters. However, contrary to the traditional robust approach, our method does not explicitly require the definition of uncertainty sets.

First, we propose a robust structural electric system model including transmission, thermic constraints and storage with auto correlated and spatially cross-correlated residual demand parameters. Then, similar to [12], we use principal component analysis in order to capture correlations between residual demand parameters and create a decorrelated vector by projection along principal components. Finally, we generate a set of bins for the training data to estimate the order statistics differential distribution, that is the conditional distribution of variations of the uncertain parameter, using parametric polynomial regression. Our method can be related to quantile regression but allows more flexibility in the usage of the training data. Section II presents the formulation of our structural electric system model. Our polynomial regression method is described in Section III. Then, our model is applied in Section IV to the case of the French region Auvergne Rhône-Alpes, with interesting results regarding flexibility requirements for electricity mixes with strong renewable penetration, before concluding in Section V.

II. FORMULATION OF THE STUCTURAL ELECRIC SYSTEM MODEL

Modern electric systems including renewable generation technologies can be broken down into a series of simple components: electric load, renewable production units, dispatchable production units with thermal limits, storage units and a transmission network. For simplicity, we consider a single region and neglect transmission constraint in the following model. Residual demand can be expressed as a linear combination of electric load and production from renewable generation units. We denote the installed capacity of renewable technology $j \in \mathcal{J}$ as $U_j \geq 0$ with uncertain capacity factor $\xi_j = \overline{\xi_j} + \widehat{\xi_j}\eta_j$, with nominal or average value noted $\overline{\xi_j}$ and standard deviation $\widehat{\xi_j}$. η_j is a random variable with zero mean and standard deviation equal to one. The set of dispatchable generation technologies with thermal limits is noted $\mathcal{G}$, with $g \in \mathcal{G}$ and installed capacity $U_g \geq 0$. Finally, we define the installed capacity of storage as $U_e \geq 0$. Renewable technologies, the dispatchable generation technologies and the storage technology respectively have unit investment costs c_j^{FOM}, c_g^{FOM} and c_e^{FOM}, with $c_j^{FOM}, c_g^{FOM}, c_e^{FOM} \geq 0$. Finally, we respectively note c_g^V and c_g^{STUP} the variable and start-up costs of dispatchable technology $g \in \mathcal{G}$.

The structural cost-minimization problem for any electricity system, neglecting spatial transfers and transmission network, can then be defined as follows:

$$\min_{\boldsymbol{U},\boldsymbol{q},\boldsymbol{z},\boldsymbol{s}} \sum_{j>1} c_j^{FOM} U_j + \sum_g c_g^{FOM} U_g + c_e^{FOM} U_e + \sum_{t\in\mathcal{T}} \sum_g c_g^V \left(q_{gt}^{UB} + q_{gt}^{LB} + q_{gt}^V\right) + c_g^{STUP} z_{gt} \qquad \mathbf{(1)}$$

Such that :

$$\overline{\xi_{1t}} + \widehat{\xi_{1t}}\eta_{1t} - \sum_{j>1} \left(\overline{\xi_{jt}} + \widehat{\xi_{jt}}\eta_{jt}\right) U_j - \sum_g q_{gt} + e_t^+ - e_t^- \leq \alpha^+ \qquad \mathbf{(2)}$$

$$-\overline{\xi_{1t}} - \widehat{\xi_{1t}}\eta_{1t} + \sum_{j>1} \left(\overline{\xi_{jt}} + \widehat{\xi_{jt}}\eta_{jt}\right) U_j + \sum_g q_{gt} - e_t^+ + e_t^- \leq \alpha^+ \qquad \mathbf{(3)}$$

$$u_{gt} - u_{gt-1} = z_{gt} - v_{gt} \qquad ,\forall g \in \mathcal{G} \quad \mathbf{(4a)}$$

$$z_{gt} + v_{gt} \leq 1 \qquad ,\forall g \in \mathcal{G} \quad \mathbf{(4b)}$$

$$\omega_{gt}^1 = q_{gt} - \omega_{gt}^2 \qquad ,\forall g \in \mathcal{G} \quad \mathbf{(5)}$$

$$\omega_{gt}^1 - \omega_{gt-1}^1 \leq \overline{r_g} \qquad ,\forall g \in \mathcal{G} \quad \mathbf{(6a)}$$

$$\omega_{gt-1}^1 - \omega_{gt}^1 \leq \underline{r_g} \qquad ,\forall g \in \mathcal{G} \quad \mathbf{(6b)}$$

$$\omega_{gt}^1 \leq U_g\left(\overline{q_g} - \underline{q_g}\right) \qquad ,\forall g \in \mathcal{G} \quad (\mathbf{7a})$$

$$\omega_{gt}^1 \leq u_{gt}\Omega \qquad ,\forall g \in \mathcal{G} \quad (\mathbf{7b})$$

$$\omega_{gt}^2 \leq U_g\underline{q_g} \qquad ,\forall g \in \mathcal{G} \quad (\mathbf{8})$$

$$\omega_{gt}^2 = U_g\underline{q_g} - (1 - u_{gt})\Omega + s_{gt} \qquad ,\forall g \in \mathcal{G} \quad (\mathbf{9})$$

$$s_{gt} \leq (1 - u_{gt})\Omega - U_g\underline{q_j} + u_{gt}\Omega \qquad ,\forall g \in \mathcal{G} \quad (\mathbf{10})$$

$$u_{gt} \geq \sum_{t'>t-M_j^U} z_{gt} \qquad ,\forall g \in \mathcal{G} \quad (\mathbf{11})$$

$$1 - u_{gt} \geq \sum_{t'>t-M_j^D} v_{gt} \qquad ,\forall g \in \mathcal{G} \quad (\mathbf{12})$$

$$e_t = e_{t-1} + \sqrt{\eta_e}e_t^+ - \frac{e_t^-}{\sqrt{\eta_e}} \qquad (\mathbf{13})$$

$$e_t \leq \bar{e}U_e \qquad (\mathbf{14})$$

$$e_t \geq \underline{e}U_e \qquad (\mathbf{15})$$

$$e_t^+ \leq l_t(\bar{e}U_e - e_t) \qquad (\mathbf{16})$$

$$e_t^- \leq (1 - l_t)(e_t - \underline{e}U_e) \qquad (\mathbf{17})$$

$$u_{gt} \in \{0,1\} \qquad ,\forall g \in \mathcal{G} \quad (\mathbf{18})$$

$$v_{gt} \in \{0,1\} \qquad ,\forall g \in \mathcal{G} \quad (\mathbf{19})$$

$$z_{gt} \in \{0,1\} \qquad ,\forall g \in \mathcal{G} \quad (\mathbf{20})$$

$$l_t \in \{0,1\} \qquad ,\forall g \in \mathcal{G} \quad (\mathbf{21})$$

Each constraint must hold for each time period $t \in \mathcal{T}$, where the set $\mathcal{T}$ is used to index time. Equation **(1)** corresponds to the sum of the investment cost of the electricity mix and short-term dispatching and start-up costs. Equations **(2)** and **(3)** constrain net generation, which is the sum of electricity generation minus electric load and storage, to lie in the interval defined by $[-\alpha^+, \alpha^+]$. e_t corresponds to the stock of electricity stored in t, while e_t^+ and e_t^- are respectively equal to the quantity of electricity stored and released in t. We note $[a]^+ =$

$\max(a, 0)$ for convenience and define $\boldsymbol{U} = (U_i)_{1 \le i \le 1+|\mathcal{J}|}$ the vector of installed capacities for residual demand components, with the first element corresponding to electric load so $U_1 = -1$. Equations **(4a)** to **(21)** together formalize as a set of linear constraints commitment state, starting-up decisions and output limits for dispatchable generators. For any time period $t \in \mathcal{T}$, ω_{gt}^2 corresponds to the minimum-production level of generator $g \in \mathcal{G}$, while ω_{gt}^1 is an auxiliary variable equal to the generation volume above minimum-production level. u_{gt}, v_{gt} and z_{gt} are all binary variables respectively corresponding to the commitment state, start-up and shut-down decision of generator $g \in \mathcal{G}$. Finally, l_t corresponds to the charging state of batteries, with $l_t = 1$ when batteries store electricity.

III. A FLEXIBLE POLYNOMIAL REGRESSION BASED DYNAMIC ROBUST MODEL FOR AUTOCORRELATED RESIDUAL DEMAND

For each t, we define the random vector $\widehat{\boldsymbol{\xi}_t} = (\widehat{\eta_{1t}}, \dots, \widehat{\eta_{nt}})^T \in \mathbb{R}^{(1\times n)}$, where n is the number of uncertain parameters. We assume that the random process $\{\widehat{\boldsymbol{\xi}_t}\}_{t\in\mathcal{T}}$ is (weakly) stationary. Using the variance-covariance matrix of $\widehat{\boldsymbol{\xi}_t}$, we use its corresponding matrix of eigenvectors $\boldsymbol{\Phi}$ so that $\boldsymbol{\mu}_t = \boldsymbol{\Phi}^T \widehat{\boldsymbol{\xi}_t}$ has a diagonal variance-covariance matrix, ie has uncorrelated components. We prove that autocovariance matrices of $\boldsymbol{\mu}_t$ are diagonal at all lag values.

III.A. Decorrelation procedure of the autocorrelated vector and formulation as vector polynomial

We start by defining the variance-covariance matrix of $\widehat{\boldsymbol{\xi}_t}$ as $\boldsymbol{\Sigma}_\xi \in \mathbb{R}^{n\times n}$. We define $\boldsymbol{\Lambda} \in \mathbb{R}^{n\times n}$ the matrix of its eigenvalues and $\boldsymbol{V} \in \mathbb{R}^{n\times n}$ the corresponding matrix of eigenvectors. Then, using the orthogonality of $\boldsymbol{V}$, we have $\boldsymbol{V}^T = \boldsymbol{V}^{-1}$ so that:

$$\boldsymbol{\Sigma}_\xi \boldsymbol{V} = \boldsymbol{V}\boldsymbol{\Lambda} \Leftrightarrow \boldsymbol{V}^T \boldsymbol{\Sigma}_\xi \boldsymbol{V} = \boldsymbol{V}^T \boldsymbol{V} \boldsymbol{\Lambda} = \boldsymbol{\Lambda} \in \mathbb{R}^{n\times n} \qquad (\mathbf{22})$$

Our transformed random vector $\boldsymbol{\mu}_t = \boldsymbol{V}^T \widehat{\boldsymbol{\xi}_t}$ has the following diagonal variance-covariance matrix :

$$\boldsymbol{\Sigma}_\mu = \boldsymbol{V}^T \boldsymbol{\Sigma}_\xi \boldsymbol{V} = \boldsymbol{\Lambda} \in \mathbb{R}^{n\times n} \qquad (\mathbf{23})$$

It is possible to show that for any lag value $p \in \mathbb{N}$, the autocovariance matrix of $\boldsymbol{\mu}_t$ of order p, noted $\boldsymbol{K}_p = \mathbb{E}(\boldsymbol{\mu}_t \boldsymbol{\mu}_{t-p})$ is diagonal. As $\boldsymbol{\Sigma}_\mu$ is diagonal, we have by construction

$\mathrm{cov}(\mu_{it}, \mu_{jt}) = 0, \forall (i,j) \in [1,n]^2, i \neq j$. Assume that $\mathrm{cov}(\mu_{it}, \mu_{i,t-p}) \neq 0$ and $\mathrm{cov}(\mu_{it}, \mu_{j,t-p}) \neq 0$. Then, there exists a couple $(\phi_{jj,p}, \phi_{ij,p}) \in (\mathbb{R}^*)^2$ such that:

$$\mu_{jt} = \phi_{jj,p}\mu_{j,t-p} + \varepsilon_{j,t-p} \quad (\mathbf{24a})$$

$$\mu_{it} = \phi_{ij,p}\mu_{j,t-p} + \varepsilon_{i,t-p} \quad (\mathbf{24b})$$

With $\mathbb{E}(\varepsilon_{j,t-p}) = \mathbb{E}(\varepsilon_{i,t-p}) = \mathbb{E}(\varepsilon_{j,t-p}\mu_{j,t-p}) = \mathbb{E}(\varepsilon_{i,t-p}\mu_{j,t-p}) = 0$. We can reformulate the above equations as follows:

$$\mu_{it} = \phi_{ij,p}\phi_{jj,p}^{-1}(\mu_{jt} - \varepsilon_{j,t-p}) + \varepsilon_{i,t-p} = \phi_{ij,p}\phi_{jj,p}^{-1}\mu_{jt} + \varepsilon'_{i,t-p} \quad (\mathbf{25})$$

Where $\varepsilon'_{i,t-p} = \varepsilon_{i,t-p} - \phi_{ij,p}\phi_{jj,p}^{-1}\varepsilon_{j,t-p}$. By assumption, $\phi_{ij,p}\phi_{jj,p}^{-1} \neq 0$ so $\mathrm{cov}(\mu_{it}, \mu_{jt}) \neq 0$. We reach a contradiction, so $\mathrm{cov}(\mu_{it}, \mu_{jt}) = 0 \Longrightarrow \mathrm{cov}(\mu_{it}, \mu_{j,t-p}) = 0$. There exists a diagonal matrix $\mathbf{\Phi} = \mathrm{diag}(\phi_{ii})_{i \le n} \in \mathbb{R}^{n \times n}$ and a vector $\boldsymbol{\epsilon}_{t-1} \in \mathbb{R}^{n \times 1}$ such that we have the following VAR(1) model:

$$\boldsymbol{\mu}_t = \mathbf{\Phi}\boldsymbol{\mu}_{t-1} + \boldsymbol{\epsilon}_{t-1} \quad (\mathbf{26})$$

Using a similar VAR model, **Lorca & Sun (2014)** build an adaptive uncertainty set where the random vector is equal to a linear combination of its previous values plus an error term, which is subjected to a budgeted uncertainty set. We take a different approach where we estimate the bounds within which the uncertain random vector can vary between successive periods. First, we start by taking the gradient of the random transformed vector $\boldsymbol{\mu}_t$:

$$\vec{\boldsymbol{\nabla}}\,\boldsymbol{\mu}_t = \vec{\boldsymbol{\nabla}}\,\mathbf{\Phi}\boldsymbol{\mu}_{t-1} + \vec{\boldsymbol{\nabla}}\,\boldsymbol{\epsilon}_{t-1} \quad (\mathbf{27})$$

Setting $\vec{\boldsymbol{\nabla}}\,\boldsymbol{\mu}_t = 0$, we have $\vec{\boldsymbol{\nabla}}\,\boldsymbol{\epsilon}_{t-1} = -\vec{\boldsymbol{\nabla}}\,\mathbf{\Phi}\boldsymbol{\mu}_{t-1}$ so that we notice that ϵ_{it-1} can implicitly be defined as a function of μ_{it-1} when μ_{it} is fixed. As $\epsilon_{it-1} = \mu_{it} - \mu_{it-1}$, we can define the variations of the uncertain transformed vector $\Delta\boldsymbol{\mu}_t = \boldsymbol{\mu}_t - \boldsymbol{\mu}_{t-1}$ as a function of $\boldsymbol{\mu}_{t-1}$. Thus, assuming $\boldsymbol{\mu}_t$ is continuous, we define a set of application φ_i^q**,** $0 \le q \le 1$, such that φ_i^q corresponds to the q-th quantile value of the random variable $\Delta\mu_{it}$ conditional on μ_{it-1}. Using polynomial regression, we can approximate the set of application φ_i^q**,** $0 \le q \le 1$ such that we have:

$$P_{n_i^q}(\mu_{it}) = \underset{P \in \mathbb{R}_{n_i^q}[\mu_{it}]}{\arg\min} \left\| P - \varphi_i^q(\mu_{it}) \right\|_{\infty,[\inf \mu_{it}, \sup \mu_{it}]} \quad (\mathbf{28})$$

There exists sets of real coefficients $\left(\mathfrak{a}_{in}^{q}\right)_{n\leq n_i^*}$ such that:

$$P_{n_i^q}(\mu_{it}) = \sum_{n=0}^{q} \mathfrak{a}_{in}^{q}\,\mu_{it}^{n} \tag{29}$$

III.B. Estimation of the polynomial approximations of the projected vector differential distribution

We estimate the set of parameters $\left(\mathfrak{a}_{in}^{q}\right)_{n\leq n_i^q}$ using a training sample of size N. In order to mitigate the expected sampling error for extreme variations of μ_i, we cut the interval $[\inf \mu_{it}, \sup \mu_{it}]$ into $|\mathcal{B}_i|$ bins of equal length. For each interval $b_i \in \mathcal{B}_i$, we note N_{b_i} the number of observations included into the bin b_i, with $\sum_{b_i} N_{b_i} = N$. For $q \in [0,1]$, we define the empirical q-th quantile value $\Delta\mu_i^{(b_i,q)}$ such that, with $\Delta\mu_{i1} \leq \cdots \leq \Delta\mu_{im} \leq \cdots \leq \Delta\mu_{iN_{b_i}}$:

$$\Delta\mu_i^{(b_i,q)} = \inf_{\substack{\Delta\mu_i \\ \Delta\mu_{i1}\leq\Delta\mu_i\leq\Delta\mu_{iN_{b_i}}}} \left\{\Delta\mu_i : \frac{1}{N_{b_i}}\left(\sum_{m=1}^{N_{b_i}} \mathbb{1}\{\Delta\mu_{im} \leq \Delta\mu_i\}\right) \geq q\right\} \tag{30}$$

By noting $\sigma_{\Delta\mu_i}^{q}$ the empirical sampling error of q-th quantile of $\Delta\mu_i$ estimated from the training sample, the expected sampling error $\sigma_{\Delta\mu_i^{(b_i,q)}}$ of any quantile $q \in [0,1]$ can be approximated as $\sigma_{\Delta\mu_i^{(b_i,q)}} \cong K(q)\sigma_{\Delta\mu_i}^{q} N_{b_i}^{-\frac{1}{2}}$, where $K(q)$ is a function satisfying $K(q) \geq 1$ and $K'(q) \geq 0$. Then, by noting M_{b_i} the median value of the bin b_i, we can estimate the following systems of equations:

$$\boldsymbol{Y}_q = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{\epsilon}_q \tag{31}$$

, where:

$$\boldsymbol{Y}_q = \begin{pmatrix} \Delta\mu_i^{(1,q)} \\ \Delta\mu_i^{(2,q)} \\ \vdots \\ \Delta\mu_i^{(|\mathcal{B}_i|,q)} \end{pmatrix}, \quad \boldsymbol{X} = \begin{pmatrix} 1 & M_1 & M_1^2 & \cdots & M_1^{n_i^*} \\ 1 & M_2 & M_2^2 & \cdots & M_2^{n_i^*} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & M_{|\mathcal{B}_i|} & M_{|\mathcal{B}_i|}^2 & \ldots & M_{|\mathcal{B}_i|}^{n_i^*} \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} \mathfrak{a}_{i0} \\ \mathfrak{a}_{i1} \\ \mathfrak{a}_{i2} \\ \vdots \\ \mathfrak{a}_{in_i^*} \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} \mathfrak{b}_{i0} \\ \mathfrak{b}_{i1} \\ \mathfrak{b}_{i2} \\ \vdots \\ \mathfrak{b}_{in_i^*} \end{pmatrix},$$

$$\boldsymbol{\epsilon}_q = \begin{pmatrix} \epsilon_{1q} \\ \epsilon_{2q} \\ \vdots \\ \epsilon_{|\mathcal{B}_i|q} \end{pmatrix}.$$

Yet, as the number of observations per bin b_i is expected to be non-constant, the MCO assumption of constant variance of the error terms is violated. Assuming the ϵ_{b_iq} are normally distributed with non-constant variance, we introduce a diagonal weighting matrix $\boldsymbol{W}$, with diagonal entry $\omega_{b_i} = \left(\sigma^2_{\Delta\mu_i^{(b_i,q)}}\right)^{-1}$. The MCO estimators are then equal to $\widehat{\boldsymbol{A}} = (\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{Y}_q$. The standard errors associated to each estimated coefficient $\widehat{\mathfrak{a}_{in}}$ are given by the following expression:

$$\sigma_{\widehat{\mathfrak{a}_{in}^q}} = \sqrt{\frac{\boldsymbol{Y}_q^T\boldsymbol{Y}_q - \widehat{\boldsymbol{A}}^T\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X}\widehat{\boldsymbol{A}}}{|\mathcal{B}_i| - n_i^q - 1}} \times \sqrt{(\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1}_{n+1}} \qquad \textbf{(32)}$$

Where $(\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1}_{n+1}$ is the $n+1$-th diagonal element of the matrix $(\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1}$. We observe that both expressions for standard error decrease with the number of bins $|\mathcal{B}_i|$. Simultaneously, increasing $|\mathcal{B}_i|$ decreases the number of observations per bins, which mechanically increases $\sigma_{\Delta\mu_i^{(b_i,q)}}$. If $\sigma_{\Delta\mu_i^{(b_i,q)}} \approx \sigma_{\Delta\mu_i^{(b_i',q)}}$ for any couple $(b_i, b_i') \in \mathcal{B}_i^2$, then all weight coefficients in matrix $\boldsymbol{W}$ decrease in similar proportions. Yet, if $\sigma_{\Delta\mu_i^{(b_i,q)}} \neq \sigma_{\Delta\mu_i^{(b_i',q)}}$, bins with a large error will have a low weight in the estimation of the polynomial coefficients and the polynomial approximation may thus be poorly reliable for forecasting limiting variations of $\Delta\mu_i$ within these bins intervals. In order to maximize the reliability of our approximation polynomials for all bins intervals, we chose the number of bins $|\mathcal{B}_i|^*$ which minimizes the variance of $N_{\mathcal{B}_i}$.

Assuming that the number of observations is higher around the mean value of $\Delta\mu_i$ (ie small variations occur more frequently than large ones), the number of observations is inversely proportional to the distance to the mean value $\overline{\Delta\mu_i}$. Under this behavioral assumption, quantile estimates corresponding to extreme values of $\Delta\mu_i$ will be associated to a higher standard error.

This feature will be captured by the prediction intervals of $\widehat{P_{n_i^q}(\mu_{it})}$. Indeed, using the propagation of uncertainty method, the standard error of $\widehat{P_{n_i^q}(\mu_{it})}$ can be expressed as:

$$\sigma_{\widehat{P_{n_i^q}(\mu_{it})}} = \sqrt{\sum_{n=0}^{n_i^q} \sigma_{\widehat{\mathrm{a}_{in}^q}}^2 \mu_{it}^{2n}} = \sqrt{\sum_{n=0}^{n_i^q} \frac{\boldsymbol{Y}_q^T \boldsymbol{Y}_q - \widehat{\boldsymbol{A}}^T \boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X} \widehat{\boldsymbol{A}}}{|\mathcal{B}_i|^* - n_i^q - 1} \times (\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X})_{n+1}^{-1} \mu_{it}^{2n}} \qquad \textbf{(33)}$$

Assuming normality of $\widehat{\boldsymbol{A}} \sim N\left(\boldsymbol{A}, \sigma_{\Delta\mu_i}^q (\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X})^{-1}\right)$ allows for exact confidence intervals and prediction intervals. Considering polynomials $\widehat{P_{n_i^q}(\mu_{it})}$ as linear combinations, the variance of the predicted values $\widehat{P_{n_i^q}(\mu_{it})}$ are respectively $\sigma_{\Delta\mu_i}^q + \boldsymbol{X}_t^T \frac{\boldsymbol{Y}_q^T \boldsymbol{Y}_q - \widehat{\boldsymbol{A}}^T \boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X} \widehat{\boldsymbol{A}}}{|\mathcal{B}_i|^* - n_i^q - 1} \times (\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X})^{-1} \boldsymbol{X}_t$ with $\boldsymbol{X}_t = \left(1, \mu_{it}, \mu_{it}^2, \dots, \mu_{it}^{n_i^q}\right)^T \in \mathbb{R}^{(n_i^q+1)\times 1}$. The prediction interval of the q-th quantile value of $\Delta\mu_i$, conditional on $\boldsymbol{X}_t$, can thus be expressed as:

$$\mathcal{U}_i(q, \alpha, |\mathcal{B}_i|^* | \boldsymbol{X}_t) = \left[\widehat{\boldsymbol{A}} \boldsymbol{X}_t \pm t_{1-\frac{\alpha}{2}; |\mathcal{B}_i|^* - n_i^q} \times \sqrt{\sigma_{\Delta\mu_i}^q + \boldsymbol{X}_t^T \frac{\boldsymbol{Y}_q^T \boldsymbol{Y}_q - \widehat{\boldsymbol{A}}^T \boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X} \widehat{\boldsymbol{A}}}{|\mathcal{B}_i|^* - n_i^q - 1} \times (\boldsymbol{X}^T \boldsymbol{W} \boldsymbol{X})^{-1} \boldsymbol{X}_t} \right] \qquad \textbf{(34)}$$

Where $t_{1-\frac{\alpha}{2}; |\mathcal{B}_i|^* - n_i^q}$ is the upper critical value of a Student distribution with $|\mathcal{B}_i|^* - n_i^q$ degrees of liberty for quantile $1 - \frac{\alpha}{2}$.

The prediction interval is wider than the confidence interval as it accounts for the fluctuations of $\Delta\mu_i$. Our prediction intervals can be interpreted as follows: if for any time period t, we observe the variation $\Delta\mu_{it}$ for a given value μ_{it-1} and repeat the experiment, the q-th quantile of the distribution of variations will be included in the prediction interval $\mathcal{U}_i(q, \alpha, |\mathcal{B}_i|^* | \boldsymbol{X}_t)$ in $1 - \alpha\%$ of the cases. Taking $\alpha = 0.05$, these is only 5% chances that $\Delta\mu_i^q \notin \mathcal{U}_i(q, \alpha, |\mathcal{B}_i|^* | \boldsymbol{X}_t)$. Assuming that the prediction error is normally distributed around the mean predicted value, it is straightforward to see that :

$$\mathbb{P}(\Delta\mu_i \leq \sup \mathcal{U}_i(q, \alpha, |\mathcal{B}_i|^* | \boldsymbol{X}_t)) = q - \frac{\alpha}{2} \qquad \textbf{(35a)}$$

$$\mathbb{P}(\Delta\mu_i \geq \inf \mathcal{U}_i(1 - q, \alpha, |\mathcal{B}_i|^* | \boldsymbol{X}_t)) = 1 - q + \frac{\alpha}{2} \qquad \textbf{(35b)}$$

So we have:

$$\mathbb{P}(\Delta\mu_i \notin [\inf \mathcal{U}_i(1-q,\alpha,|\mathcal{B}_i|^*|\boldsymbol{X}_t)\,;\sup \mathcal{U}_i(q,\alpha,|\mathcal{B}_i|^*|\boldsymbol{X}_t)]) = 2-2q+\alpha \qquad (\mathbf{35c})$$

Finally, we can define the empirical approximation of φ_i^q as:

$$\widehat{\varphi_i^q}(\alpha,|\mathcal{B}_i|^*|\boldsymbol{X}_t) = \widehat{\boldsymbol{A}}\boldsymbol{X}_0 + t_{1-\frac{\alpha}{2};|\mathcal{B}_i|^*-n_i^*} \times \sqrt{\sigma_{\Delta\mu_i}^q + \boldsymbol{X}_0^T\,\frac{\boldsymbol{Y}_q^T\boldsymbol{Y}_q - \widehat{\boldsymbol{A}}^T\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X}\widehat{\boldsymbol{A}}}{|\mathcal{B}_i|^* - n_i^* - 1} \times (\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}_0} \qquad (\mathbf{36})$$

Our method actually mimics a polynomial quantile regression, as we model the conditional dependence of quantiles of the endogenous variable on a polynomial. However, as the density function of the dependent variable is expected to be higher around its mean value, conditional quantiles of $\Delta\mu_i$ associated to extreme values of μ_i are expected to be highly noisy. This would likely bias the conditional quantile regression estimator.

Our binning strategy allows us to decrease the overall variance of sample conditional quantiles of $\Delta\mu_i$ first. Then, applying weighting reduces the influence of bins with high conditional quantile variance on the estimator value. Finally, the larger prediction interval associated to a bin with lower weight accounts for its smaller contribution in the estimation procedure and the higher probability of forecasting error for values of the exogenous variable included in the bin. Though parametric, our method allows us to account for differences in the density of observations, especially for extreme values, so we provide robust predictions of the conditional interval within which our unknown parameter can vary.

III.C. Dynamic robust reformulation of the optimization model

Applying the above results, the optimization model becomes:

$$\min_{\boldsymbol{U},\boldsymbol{q},\boldsymbol{z},\boldsymbol{s}} \sum_{j>1} c_j^{FOM} U_j + \sum_g c_g^{FOM} U_g + c_e^{FOM} U_e + \sum_{t\in\mathcal{T}}\sum_g c_g\left(q_{gt}^H + q_{gt}^L + q_{gt}^V\right) + c_g^{STUP} z_{gt} \qquad (\mathbf{37})$$

such that **(4a)-(20)** hold and :

$$\bar{\xi}_{1t} + \widehat{\xi}_1\pi_{1t}^+ - \sum_{j>1}\left(\bar{\xi}_{jt} + \widehat{\xi_{jt}}\pi_{jt}^+\right)U_j - \sum_g q_{gt}^H + \Delta e_t^{+H} - e_t^{-H} \le \alpha^+ \qquad (\mathbf{38a})$$

$$-\bar{\xi}_{1t} - \widehat{\xi}_1\pi_{1t}^+ + \sum_{j>1}\left(\bar{\xi}_{jt} + \widehat{\xi_{jt}}\pi_{jt}^+\right)U_j + \sum_g q_{gt}^H - \Delta e_t^{+H} + e_t^{-H} \le \alpha^+ \qquad (\mathbf{38b})$$

$$\bar{\xi}_{1t} + \widehat{\xi}_1 \pi_{1t}^- - \sum_{j>1} (\bar{\xi}_{jt} + \widehat{\xi_{jt}} \pi_{jt}^-) U_j - \sum_g q_{gt}^L + \Delta e_t^{+L} - e_t^{-L} \leq \alpha^+ \quad (\mathbf{39a})$$

$$-\bar{\xi}_{1t} - \widehat{\xi}_1 \pi_{1t}^- + \sum_{j>1} (\bar{\xi}_{jt} + \widehat{\xi_{jt}} \pi_{jt}^-) U_j + \sum_g q_{gt}^L - \Delta e_t^{+L} + e_t^{-LB} \leq \alpha^+ \quad (\mathbf{39b})$$

$$\bar{\xi}_{1t} + \widehat{\xi}_1 \pi_{1t}^V - \sum_{j>1} (\bar{\xi}_{jt} + \widehat{\xi_{jt}} \pi_{jt}^V) U_j - \sum_g q_{gt}^V + \Delta e_t^{+V} - e_t^{-V} \leq \alpha^+ \quad (\mathbf{40a})$$

$$-\bar{\xi}_{1t} - \widehat{\xi}_1 \pi_{1t}^V + \sum_{j>1} (\bar{\xi}_{jt} + \widehat{\xi_{jt}} \pi_{jt}^V) U_j + \sum_g q_{gt}^V - \Delta e_t^{+V} + e_t^{-V} \leq \alpha^+ \quad (\mathbf{40b})$$

$$\boldsymbol{\pi}_t^+ = \boldsymbol{\pi}_{t-1}^+ + \boldsymbol{V}^{-T} \widehat{\boldsymbol{\varphi}^q} (\alpha, |\boldsymbol{\mathcal{B}}|^* | \boldsymbol{V}^T \boldsymbol{\pi}_{t-1}^+) \quad (\mathbf{41a})$$

$$\boldsymbol{\pi}_t^- = \boldsymbol{\pi}_{t-1}^- + \boldsymbol{V}^{-T} \widehat{\boldsymbol{\varphi}^{1-q}} (\alpha, |\boldsymbol{\mathcal{B}}|^* | \boldsymbol{V}^T \boldsymbol{\pi}_{t-1}^-) \quad (\mathbf{41b})$$

$$\boldsymbol{\pi}_t^V = \boldsymbol{\pi}_{t-1}^V + \left(\mathbf{1}^{(n\times 1)} - \boldsymbol{w_t}\right) \circ \left(\boldsymbol{V}^{-T} \widehat{\boldsymbol{\varphi}^q} (\alpha, |\boldsymbol{\mathcal{B}}|^* | \boldsymbol{V}^T \boldsymbol{\pi}_{t-1}^V)\right) + \boldsymbol{w}_t \circ \left(\boldsymbol{V}^{-T} \widehat{\boldsymbol{\varphi}^{1-q}} (\alpha, |\boldsymbol{\mathcal{B}}|^* | \boldsymbol{V}^T \boldsymbol{\pi}_{t-1}^V)\right) \quad (\mathbf{41c})$$

$$\begin{aligned} -\Delta \bar{\boldsymbol{\xi}}_t \circ \boldsymbol{U} - \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T} \widehat{\boldsymbol{\varphi}^q} (\alpha, |\boldsymbol{\mathcal{B}}|^* | \boldsymbol{V}^T \boldsymbol{\pi}_{t-1}^V)\right) \circ \boldsymbol{U} \\ \leq \Delta \bar{\boldsymbol{\xi}}_t \circ \boldsymbol{U} + \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T} \widehat{\boldsymbol{\varphi}^{1-q}} (\alpha, |\boldsymbol{\mathcal{B}}|^* | \boldsymbol{V}^T \boldsymbol{\pi}_{t-1}^V)\right) \circ \boldsymbol{U} + \boldsymbol{M} \circ \left(\mathbf{1}^{(n\times 1)} - \boldsymbol{w}_t\right) \end{aligned} \quad (\mathbf{42a})$$

$$\begin{aligned} -\Delta \bar{\boldsymbol{\xi}}_t \circ \boldsymbol{U} - \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T} \widehat{\boldsymbol{\varphi}^q} (\alpha, |\boldsymbol{\mathcal{B}}|^* | \boldsymbol{V}^T \boldsymbol{\pi}_{t-1}^V)\right) \circ \boldsymbol{U} \\ \geq \Delta \bar{\boldsymbol{\xi}}_t \circ \boldsymbol{U} + \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T} \widehat{\boldsymbol{\varphi}^{1-q}} (\alpha, |\boldsymbol{\mathcal{B}}|^* | \boldsymbol{V}^T \boldsymbol{\pi}_{t-1}^V)\right) \circ \boldsymbol{U} - \boldsymbol{M} \circ \boldsymbol{w}_t \end{aligned} \quad (\mathbf{42b})$$

$$\boldsymbol{w}_t \in [0\,;1]^{(n\times 1)} \quad (\mathbf{43})$$

$$\boldsymbol{V}^T \boldsymbol{\pi}_0^+ \geq \boldsymbol{\Psi}^+ \quad (\mathbf{44a})$$

$$\boldsymbol{V}^T \boldsymbol{\pi}_0^- \leq \boldsymbol{\Psi}^- \quad (\mathbf{44b})$$

$$\boldsymbol{V}^T \boldsymbol{\pi}_0^V = \boldsymbol{\Psi}^V \quad (\mathbf{44c})$$

As there may exist no single set of dispatching decisions which simultaneously verify **(38a)** to **(40c)**, we introduce superscripts L, H and V which respectively correspond to subset of dispatching decisions associated to the lower, upper and most volatile conditional trajectories of residual demand. Constraints **(4a)-(20)** hold for each subset of dispatching decision variables. Contrary to the methodology adopted in Chapter 1, the extreme trajectories are not computed *a priori* over the optimization period $\mathcal{T}$. In this case, we derive for each point of time the values of the uncertain vector that maximize residual demand conditional on the previous value of the uncertain vector. The same is done in order to compute the variations of the uncertain vector than minimize residual value and maximize its absolute variation. As each variation can be associated to a positive probability, it is possible to compute the joint probability of any trajectory of the uncertain vector. We may further introduce a probabilistic threshold so that we restrict ourselves to the worst-case residual demand trajectories which satisfy it. We propose in Appendix a linear programming approach in order to derive the worst-case set of trajectories with a probability of occurrence superior or equal to a given threshold. The variability-maximizing trajectory of uncertain parameters is computed using a quadratic formulation, which can easily be reformulated into an equivalent linear form.

(41a) and **(41b)** model the dynamics of the uncertain positive and negative projected vectors $\boldsymbol{\pi}_t^+$ and $\boldsymbol{\pi}_t^-$, where $\boldsymbol{\pi}_t^+ - \boldsymbol{\pi}_{t-1}^+$ corresponds to the maximum positive variation conditional on $\boldsymbol{\pi}_{t-1}^+$ with probability inferior or equal to $q - \frac{\alpha}{2}$. This means that, for each time period, there is a probability inferior to $1 - q + \frac{\alpha}{2}$ that the observed variations of the uncertain parameters are larger than $\boldsymbol{\pi}_t^+ - \boldsymbol{\pi}_{t-1}^+$. The same reasoning applies to the largest negative variation of the uncertain parameters. **(41C)** is a weighted sum of the maximum and minimum variations possible given the couple (q, α), where $w_{it} = 1$ if $\widehat{\boldsymbol{\varphi}^q}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V)$ is superior or equal to the absolute value of $\widehat{\boldsymbol{\varphi}^{1-q}}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V)$. Thus, our model chooses the variation direction such that for each parameters, the total absolute variation of the its nominal and uncertain parts is maximized. By noting that, for any pair $(a, b) \in \mathbb{R}^2$, $a \geq |b| \Leftrightarrow -a \leq b \leq a$, we have the term by term system of inequalities:

$$\Delta\bar{\boldsymbol{\xi}}_t \circ \boldsymbol{U} + \left(\hat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^q}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^+)\right) \circ \boldsymbol{U} \geq \left|\Delta\bar{\boldsymbol{\xi}}_t \circ \boldsymbol{U} + \left(\hat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^{1-q}}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^-)\right) \circ \boldsymbol{U}\right| \qquad \mathbf{(45)}$$

$$\Leftrightarrow \begin{cases} -\Delta\bar{\boldsymbol{\xi}}_t \circ \boldsymbol{U} - \left(\hat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^q}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right) \circ \boldsymbol{U} \leq \Delta\bar{\boldsymbol{\xi}}_t \circ \boldsymbol{U} + \left(\hat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^{-}}\left(\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V; q, \alpha, |\boldsymbol{\mathcal{B}}|^*\right)\right) \circ \boldsymbol{U} \\ \Delta\bar{\boldsymbol{\xi}}_t \circ \boldsymbol{U} + \left(\hat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^q}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right) \circ \boldsymbol{U} \geq \Delta\bar{\boldsymbol{\xi}}_t \circ \boldsymbol{U} + \left(\hat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^{1-q}}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right) \circ \boldsymbol{U} \end{cases}$$

The operator $\circ$ corresponds to the term by term or Hadamard product. By definition, we have $\widehat{\boldsymbol{\varphi}^q}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V) \geq \widehat{\boldsymbol{\varphi}^{1-q}}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V)$ so we are only interested in the first inequality. Then, by defining the vector $\boldsymbol{w}_t \in [0\,;1]^{(n\times1)}$, it can be reformulated as the following equivalent system of constrains:

$$\begin{cases} -\Delta\bar{\boldsymbol{\xi}}_t \circ \boldsymbol{U} - \left(\hat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^q}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right) \circ \boldsymbol{U} \leq \Delta\bar{\boldsymbol{\xi}}_t \circ \boldsymbol{U} + \left(\hat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^{1-q}}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right) \circ \boldsymbol{U} + \boldsymbol{M} \circ \left(\mathbf{1}^{(n\times1)} - \boldsymbol{w}_t\right) \\ -\Delta\bar{\boldsymbol{\xi}}_t \circ \boldsymbol{U} - \left(\hat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^q}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right) \circ \boldsymbol{U} \geq \Delta\bar{\boldsymbol{\xi}}_t \circ \boldsymbol{U} + \left(\hat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^{1-q}}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right) \circ \boldsymbol{U} - \boldsymbol{M} \circ \boldsymbol{w}_t \end{cases}$$

We use the “big-M” formulation where $\boldsymbol{M} \in \mathbb{R}^{n\times1}$ and $\boldsymbol{M} \gg \mathbf{0}^{n\times1}$. However, this formulation maximizes the variability of the uncertain parameters separately. There is no theoretical guarantee that it simultaneously maximizes the total variability of residual demand. Yet, by noticing that $\forall i \leq n, \widehat{\varphi_i^q}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V) \geq \widehat{\varphi_i^{1-q}}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V)$, we can deduce that $\widehat{\boldsymbol{\varphi}^q}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_t^V)^T\boldsymbol{U}$ corresponds to the largest positive variation of residual demand in t. Similarly, $\widehat{\boldsymbol{\varphi}^{1-q}}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_t^V)^T\boldsymbol{U}$ yields the largest negative variation of residual demand. Thus, we are reduced to only comparing the absolute values of two terms. Again, as we are interested in the total absolute variation of the residual demand, we must account for the average variation of residual demand. As previously, we can reformulate the resulting inequality in order to evacuate the absolute term:

$$\left|\Delta\bar{\boldsymbol{\xi}}_t^T\boldsymbol{U} + \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^q}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V)\right)^T\boldsymbol{U}\right| \geq \left|\Delta\bar{\boldsymbol{\xi}}_t^T\boldsymbol{U} + \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^{1-q}}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V)\right)^T\boldsymbol{U}\right| \qquad (46)$$

$$\Leftrightarrow \begin{cases} -\Delta\bar{\boldsymbol{\xi}}_t^T\boldsymbol{U} - \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^q}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right)\right)^T\boldsymbol{U} \leq \Delta\bar{\boldsymbol{\xi}}_t^T\boldsymbol{U} + \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^{1-q}}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right)\right)^T\boldsymbol{U} \\ \Delta\bar{\boldsymbol{\xi}}_t^T\boldsymbol{U} + \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^q}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right)\right)^T\boldsymbol{U} \geq \Delta\bar{\boldsymbol{\xi}}_t^T\boldsymbol{U} + \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^{1-q}}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right)\right)^T\boldsymbol{U} \end{cases}$$

By introducing the binary variable w_t, we have $w_t = 1$ when **(46)** holds true with:

$$\begin{cases} -\Delta\bar{\boldsymbol{\xi}}_t^T\boldsymbol{U} - \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^q}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right)\right)^T\boldsymbol{U} \leq \Delta\bar{\boldsymbol{\xi}}_t^T\boldsymbol{U} + \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^{1-q}}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right)\right)^T\boldsymbol{U} + M(1-w_t) \\ -\Delta\bar{\boldsymbol{\xi}}_t^T\boldsymbol{U} - \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^q}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right)\right)^T\boldsymbol{U} \geq \Delta\bar{\boldsymbol{\xi}}_t^T\boldsymbol{U} + \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^{1-q}}\left(\alpha, |\boldsymbol{\mathcal{B}}|^*\left|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V\right)\right)\right)^T\boldsymbol{U} - Mw_t \end{cases}$$

Thus, we can replace equations **(40a)**,**(40b)**, **(40c)** and **(41c)**-**(43)** by the following subset of constraints:

$$\bar{\xi}_{1t} + \sum_{j>1}\bar{\xi}_{jt}U_j + \vartheta_t - \sum_g q_{gt}^V + \Delta e_t^V \leq \alpha^+ \qquad (47a)$$

$$-\bar{\xi}_{1t} - \sum_{j>1}\bar{\xi}_{jt}U_j - \vartheta_t + \sum_g q_{gt}^V - \Delta e_t^V \leq \alpha^+ \qquad (47b)$$

$$\vartheta_t = (1-\omega_t)\left(\widehat{\boldsymbol{\xi}}_t \circ \left(\boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^q}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V)\right)\right)^T\boldsymbol{U} + \omega_t\left(\widehat{\boldsymbol{\xi}}_t \circ \left(\boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^{1-q}}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V)\right)\right)^T\boldsymbol{U} \qquad (48)$$

$$\boldsymbol{\pi}_t^V = \boldsymbol{\pi}_{t-1}^V + \left(\mathbf{1}^{(n\times1)} - \boldsymbol{\omega}_t\right) \circ \left(\boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^q}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V)\right) + \boldsymbol{\omega}_t \circ \left(\boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^{1-q}}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V)\right) \qquad (49)$$

$$\begin{aligned} -\Delta\bar{\boldsymbol{\xi}}_t^T\boldsymbol{U} - \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^q}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V)\right)^T\boldsymbol{U} & \\ \leq \Delta\bar{\boldsymbol{\xi}}_t^T\boldsymbol{U} + \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^{1-q}}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V)\right)^T\boldsymbol{U} + M(1-w_t) & \qquad (50a) \end{aligned}$$

$$\begin{aligned} -\Delta\bar{\boldsymbol{\xi}}_t^T\boldsymbol{U} - \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^q}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V)\right)^T\boldsymbol{U} & \\ \geq \Delta\bar{\boldsymbol{\xi}}_t^T\boldsymbol{U} + \left(\widehat{\boldsymbol{\xi}}_t \circ \boldsymbol{V}^{-T}\widehat{\boldsymbol{\varphi}^{1-q}}(\alpha, |\boldsymbol{\mathcal{B}}|^*|\boldsymbol{V}^T\boldsymbol{\pi}_{t-1}^V)\right)^T\boldsymbol{U} - Mw_t & \qquad (50b) \end{aligned}$$

$$w_t \in [0;1] \qquad (51)$$

Where $\boldsymbol{w_t} = (w_t, w_t, \dots, w_t)^T \in \mathbb{R}^{1\times n}$**.** Finally, equations **(39a)-(39c)** control the initial values of the uncertain parameters for each separate worst case trajectories. As $\boldsymbol{V^T}$ is a matrix of parameters, we may equivalently formulate a set of constraints controlling the initial values taken by the projected parameters. $\boldsymbol{\Psi^+}$, $\boldsymbol{\Psi^-}$ and $\boldsymbol{\Psi^V}$ are adjustable positive vectors parameters included in $\mathbb{R}^{n\times 1}$ bounding the initial vectors.

IV. EMPIRICAL APPLICATION TO THE CASE OF REGION AUVERGNE RHONE-ALPES

IV.A. Estimation of the model parameters

In the following section, we present the results of the model described above applied in the case of the French region Auvergne Rhône-Alpes. This administrative region is located in the South-East of France and enjoys strong solar irradiation compared to the national average. In addition, it accounts for roughly 11.6% of French GDP, while its mean share of national electricity load equals to 13.8%.

We assume no initial generation capacities in order to better disentangle the impact of each worst-case trajectories on investment decisions. For convenience, we further assume that no investment occurs in hydroelectric production. Moreover, we constrain the variable for capacity investment to take only discrete values for nuclear, gas turbines (GT) and combined cycle gas turbine plants (CCG), and continuous values for other production and storage technologies. Nuclear investment is performed by blocks of 1.6 GW, corresponding to the rated power of the EPR Flamanville plant (the most recent nuclear power project in France), while CCG investments are made by blocs of 0.45 GW, which corresponds to the average nominal power of General Electric's 9HA.01/.02 gas turbine. Finally, GT investments are performed by blocks of 0.3 GW. Flexibility and cost assumptions for generation units can be found in **Table 1.A** and **Table 1.B.** in Appendix. We set the price of CO2 to 50€/t. We run our model for one year, so we replace investment costs in **(37)** by annuities in order to account for the life-duration of various technologies. We use a discount factor of 5%.

When considering the technical characteristics of generation technologies presented in **Table 1.A.**, one actually notices that ramping rates are high enough for each online generation units to very entirely between their minimum and maximum generation level. Indeed, as the model is defined with an hourly time resolution, nuclear plants may vary their production by up to

100% of their rated power between successive time periods. Thus, ramping constraints are actually not binding.

Using the same RTE database on the electricity consumption, solar and wind power capacity factors on the period 2013-2018, we start by normalizing each variable by subtracting its mean and dividing by normal deviation, both defined at the hourly level. The normalized variables can be interpreted as the number of standard deviations by which the original variable deviates from its mean. We restrict ourselves to the winter season as it exhibits the highest values of residual demand. Then, we use the eigenvector matrix of the variance-covariance matrix of residual demand components in order to obtain a set of three decorrelated variables.

IV.A.1. Performance of the proposed polynomial approximation method

Using this new set of projected variables, we estimate the optimal number of bins for each residual demand component. **Figures 1.A., 1.B.** and **1.C.** plot the number of bins against the variance of the number of observations per bins for electricity demand, photovoltaic and wind power data respectively.

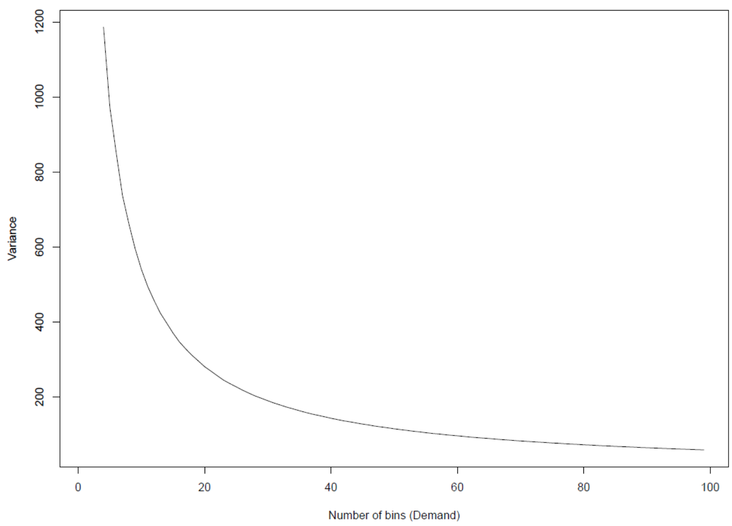

Figure 1.A.: Variance of the number of per bin observations against the number of bins (Demand)

In all three figures, the variance of the number of per bin observations is a strictly decreasing function of the number of bins. Unsurprisingly, as we define bins of equal length, the average number of per bins observations decreases. Yet, we can notice from **Figure 1.A.** that the variance of N_{b_i} for demand exhibits a steeper decrease that in the case of photovoltaic and wind power.

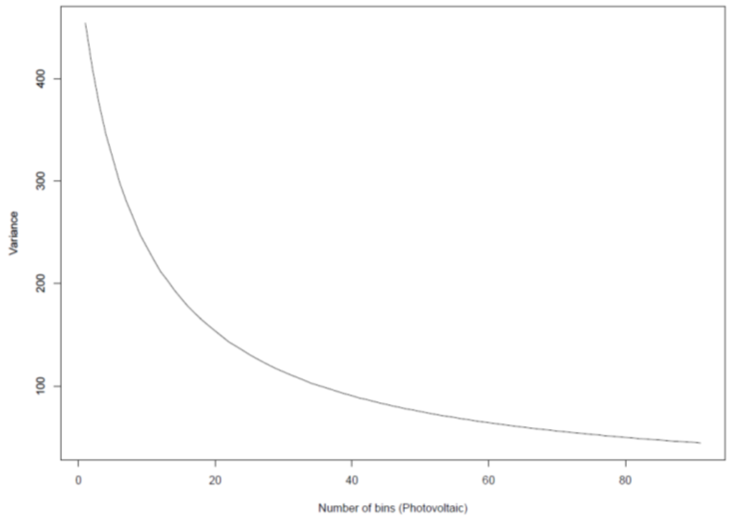

Figure 1.B.: Variance of the number of per bin observations against the number of bins (Photovoltaic power)

Recalling the inverse relationship between the number of bins and the standard deviation of our polynomial regression estimators, we chose a number of bins such that increasing it by one unit would not significantly decrease the variance of N_{b_i}. This ensures that increasing the number of bins would most likely increase the standard error of our estimators without improving the average accuracy of the empirical quantile estimate within each bin. Setting $|\mathcal{J}| = 3$ with $i = 1$ for electricity load, $i = 2$ for photovoltaic power and $i = 3$ for wind power, we chose $|\mathcal{B}_1| = 60$ and $|\mathcal{B}_2| = |\mathcal{B}_3| = 50$.

Then, for each $i \in \mathcal{J}$ and $b_i \in \mathcal{B}_i$, we estimate the empirical q-th quantile of $\Delta\mu_i$, in addition to its standard error and empirical median value of μ_i, for $q = 0.95$. Then, we estimate the equations **(25a)** and **(25b)**, setting n_i^* between 3 and 4 for all residual demand components. We finally compute the prediction interval associated to each set of estimators resulting from our polynomial regressions, setting $\alpha = 0.05$. **Figures 2.A., 2.B.** and **2.C.** respectively plot the polynomial fit corresponding to the q-th and $1 - q$-th quantiles, with prediction interval, for demand, photovoltaic and wind capacity factors.

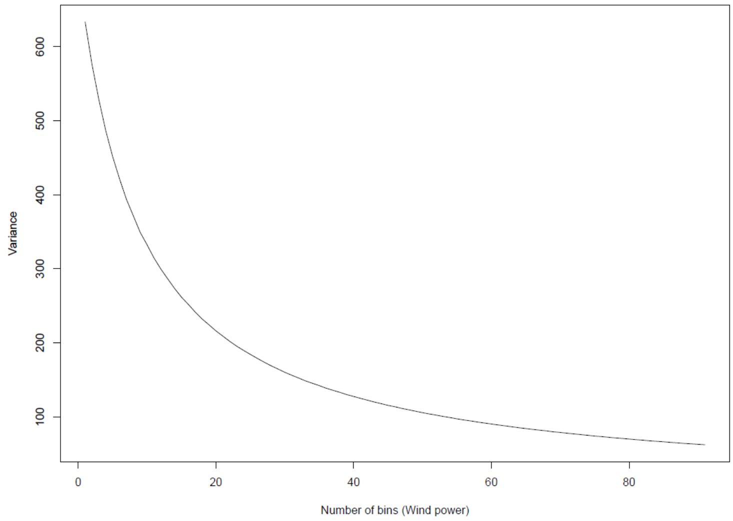

Figure 1.C.: Variance of the number of per bin observations against the number of bins (Wind power)

As observed in **Figure 2.A.**, a mean-reversing tendency for strongly negative values of the projected uncertain deviation of demand. This tendency is robust as the prediction intervals of both polynomial fitting curves are strictly positive. On the contrary, the convergence of the polynomial fitting curves for q-h and $1-q$-th quantiles for high positive values indicates the possibility of a sustained regime of high demand, as there is a probability superior or equal to q that $\Delta\mu_1$ is positive for μ_1 around 3.

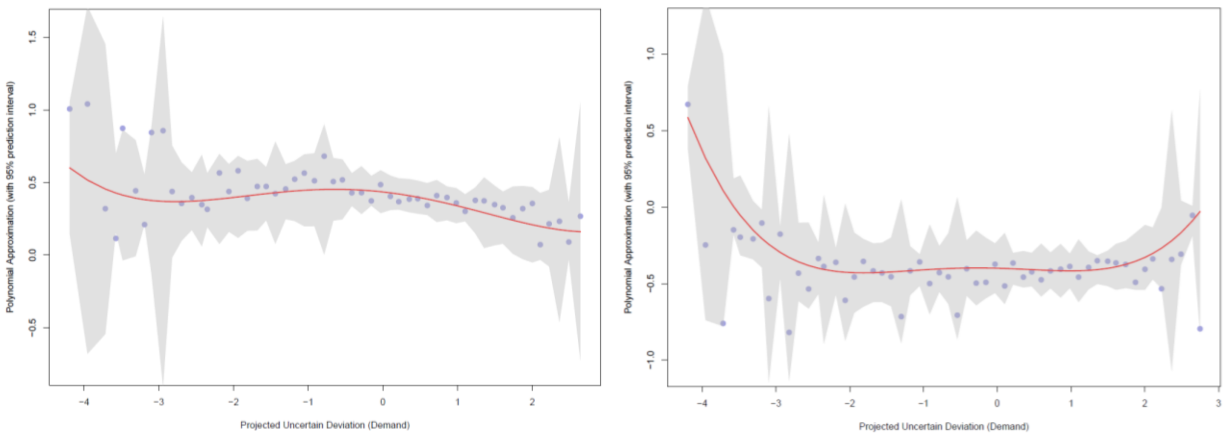

Figure 2.A.: Upper and lower quantile polynomial approximation for electricity demand

Note: The blue dots correspond to the empirical q-th quantile (left figure) and $1-q$-th quantile (right figure) associated to the median empirical value of each bin. It can be read as follows: for a projected uncertain deviation of demand equal to 0, $\Delta\mu_1 \in [-0.35, 0.5]$ approximately with a probability $2q - 1 - \alpha$.

Similar observations can be made for **Figures 2.B.** and **2.C.**. A mean-reversing behavior can be observed for extremely low levels of photovoltaic capacity factor, yet not straightforward conclusions can be drawn as regards to the sign of $\Delta\mu_2$ when μ_2 increases. This tendency is against robust as both prediction intervals are positive for extreme negative values of μ_2. Wind capacity factor exhibits a slightly different behavior. While the prediction intervals are both strictly positive for extreme left-tail values, the probability of observing a sustained regime of extremely high wind capacity factor, corresponding to the right tail of μ_3, is quasi null. Indeed, both prediction intervals are negative for μ_3 superior or equal to 2.5, which implies $\Delta\mu_3$ must be decreasing. This suggests extreme values of wind capacity factor may only occur as temporary spikes and not as a stable state.

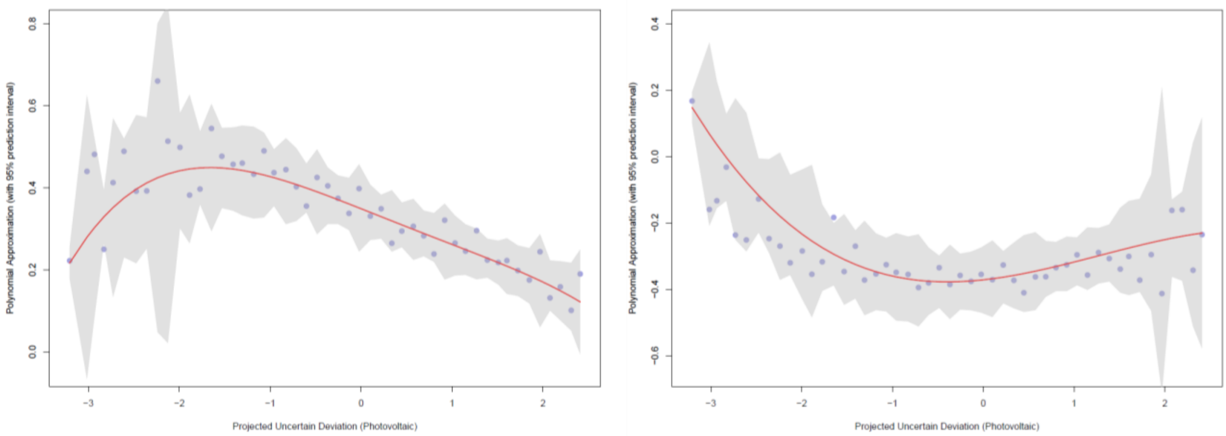

Figure 2.B.: Upper and lower quantile polynomial approximation for photovoltaic capacity factor

It may finally be noticed that our prediction intervals behave as expected: their width generally increases with the distance with respect to the median projected uncertain deviation, which translates the higher sampling error for extreme bins as they contain on average less observations than bins closer to the median. Moreover, the average width of prediction intervals decreases when estimating polynomial approximations of variations for lower values of q. **Figures 3.A., 3.B.** and **3.C.** in Appendix respectively plot the polynomial fit corresponding to the q-th and $1-q$-th quantiles for demand, photovoltaic and wind capacity factors, with $q = 0.75$. For all residual demand components, we observe a net decrease in the width of prediction interval in the neighborhood of extreme values. Generally speaking, the fact that almost all empirical quantile values are included into prediction intervals is an indicator of good performance.

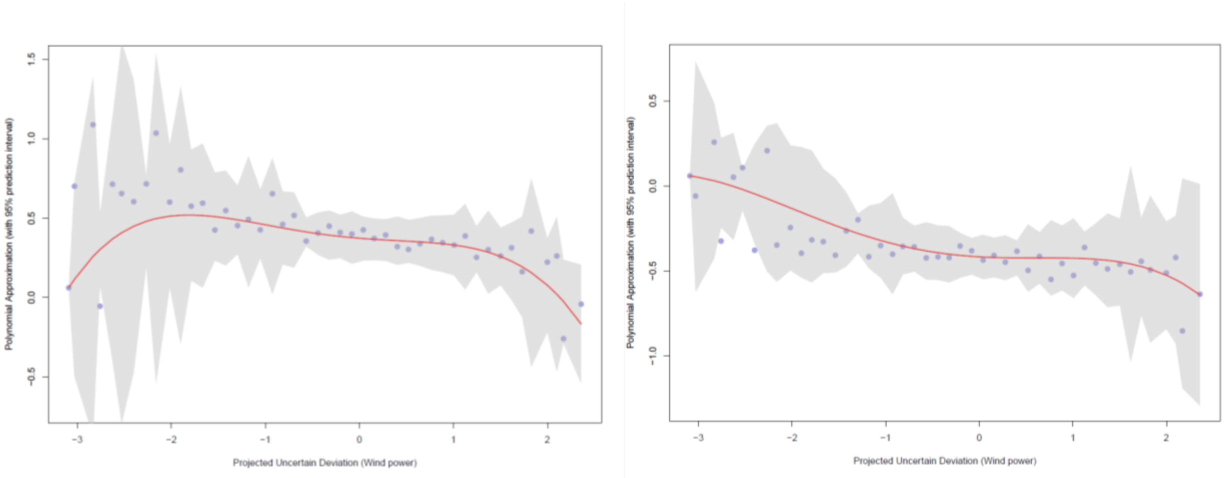

Figure 2.C.: Upper and lower quantile polynomial approximation for wind capacity factor

IV.A.2. Convergence regimes of worst-case residual demand trajectories

We now turn to the study of the dynamics of equations **(36a)** and **(36b)**. As polynomial regression provides a smooth approximation curve of $\Delta\mu_i$, $\forall i \in \mathcal{J}$, we expect $\boldsymbol{\pi}_t^+$ **and** $\boldsymbol{\pi}_t^-$ to converge to a steady-state vector. $\forall i \in \mathcal{J}$, if there exists $\pi_i^+ \in]\inf \mu_i, \sup \mu_i[$ such that $\widehat{\varphi_i^+}(\boldsymbol{V}_i^T \pi_{it-1}^+; q, \alpha, |\mathcal{B}_i|^*) = 0$, then $\lim_{t\to+\infty} \pi_{it}^+ = \pi_i^+$. The same holds for π_i^-. However, if there exists no such value, then $\lim_{t\to+\infty} \pi_{it}^+ = \sup \mu_i$ and $\lim_{t\to+\infty} \pi_{it}^- = \inf \mu_i$.

Figures 4.A., 4.B. and **4.C.** respectively plot the trajectories generated by **(36a)** and **(36b)** corresponding to various starting values for projected demand, photovoltaic and wind capacity factors, setting $q = 0.95$.[1] We first note that for all projected variables, the

---

[1] The software R does not allow us to predict future values of $\widehat{\varphi_i^+}(\boldsymbol{V}_i^T \pi_{it-1}^+; q, \alpha, |\mathcal{B}_i|^*)$ for other values than those included in the training sample. In order to circumvent this issue, we observe than $\exists \lambda \in [0,1]$, $\pi_i^+ = \lambda M_{\pi_i^+}^+ + (1-\lambda) M_{\pi_i^+}^-$, where $M_{\pi_i^+}^+$ and $M_{\pi_i^+}^-$ are the nearest bin median values such that $M_{\pi_i^+}^- \leq \pi_i^+ \leq M_{\pi_i^+}^+$. We impose $\lambda = 1$ if $\pi_i^+ > \sup M_{b_i}$ and $\lambda = 0$ if $\pi_i^+ < \inf M_{b_i}$. Then, by linearity of the prediction interval, we have

trajectories converge to an unique trajectory whatever the starting value. Indeed, we observe from

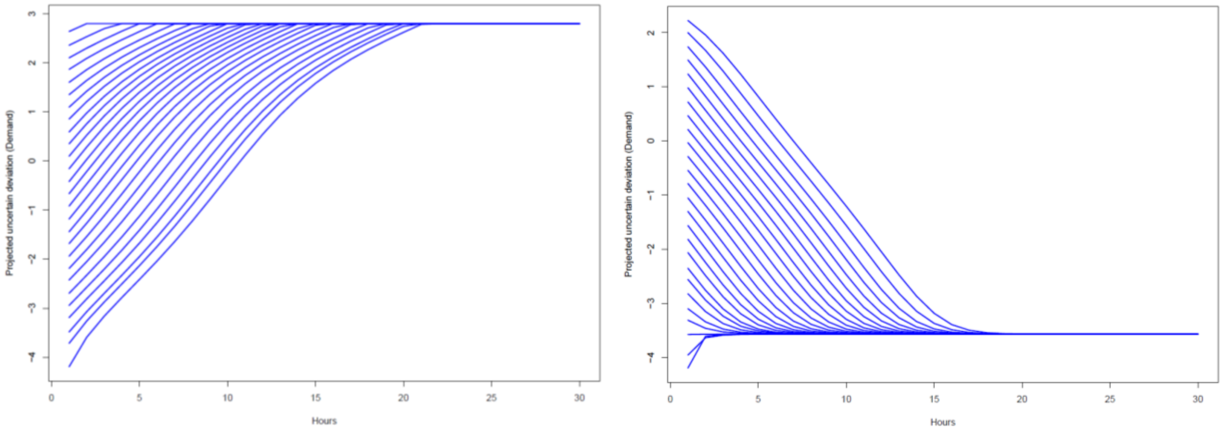

Figure 4.A.: q worst-case convergence patterns for π_{1t}^{+} and π_{1t}^{-}

Note: The left figure (resp. right figure) corresponds to the demand-maximizing (resp. demand minimizing) trajectory when to $\Delta\mu_1$ is equal to the q-th (resp. $1-q$-th) empirical quantile of its conditional distribution.

Figures 4.A. and **4.B.** that all trajectories are non-decreasing and convergence towards the upper trajectory within less than 20 hours. Interestingly, the trajectories generated by equation **(36b)** from extremely low starting values increase, while higher starting values generate non-increasing trajectories. This implies that there exist no stable trajectories such that the projected demand and solar capacity factor parameters remain at their minimum observable level.

$\widehat{\varphi_i^{+}}(\boldsymbol{V}_i^T\pi_{it-1}^{+}; q, \alpha, |\mathcal{B}_i|^*) = \lambda\widehat{\varphi_i^{+}}\left(\boldsymbol{V}_i^T M_{\pi_i^+}^{+}; q, \alpha, |\mathcal{B}_i|^*\right) + (1-\lambda)\widehat{\varphi_i^{+}}\left(\boldsymbol{V}_i^T M_{\pi_i^+}^{-}; q, \alpha, |\mathcal{B}_i|^*\right)$. We apply the same approximation method for $\widehat{\varphi_i^{-}}(\boldsymbol{V}_i^T\pi_{it-1}^{-}; q, \alpha, |\mathcal{B}_i|^*)$.

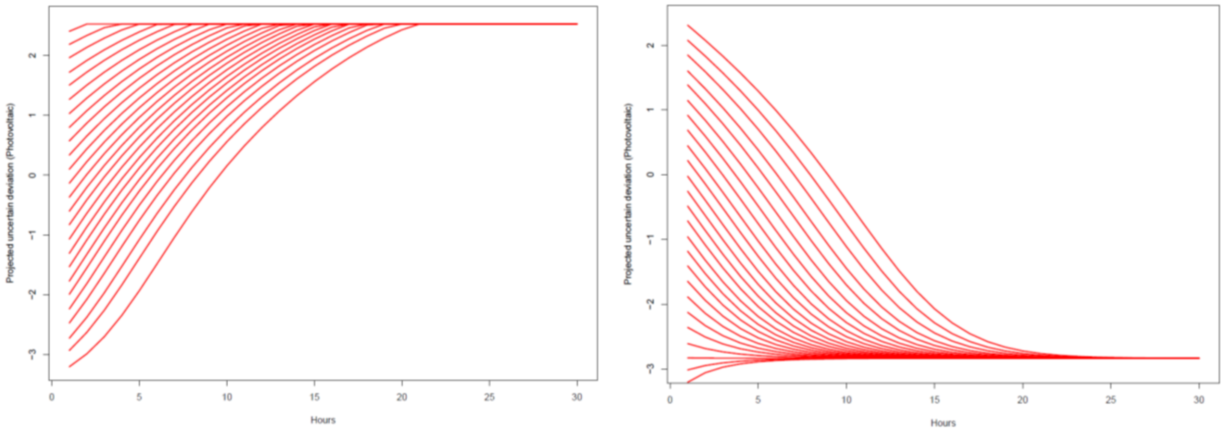

Figure 4.B.: q worst-case convergence patterns for π_{2t}^{+} and π_{2t}^{-}

Wind capacity factor exhibits a slightly different behavior. We observe from **Figure 4.C.** that **(36a)** generates decreasing trajectories for the highest starting values, while **(37b)** generates increasing trajectories for its lowest starting values. This implies that wind capacity factor converges towards worst-case trajectories with less extreme values that the initial values. Our approach is thus likely to generate less conservative solutions than a static robust methodology.

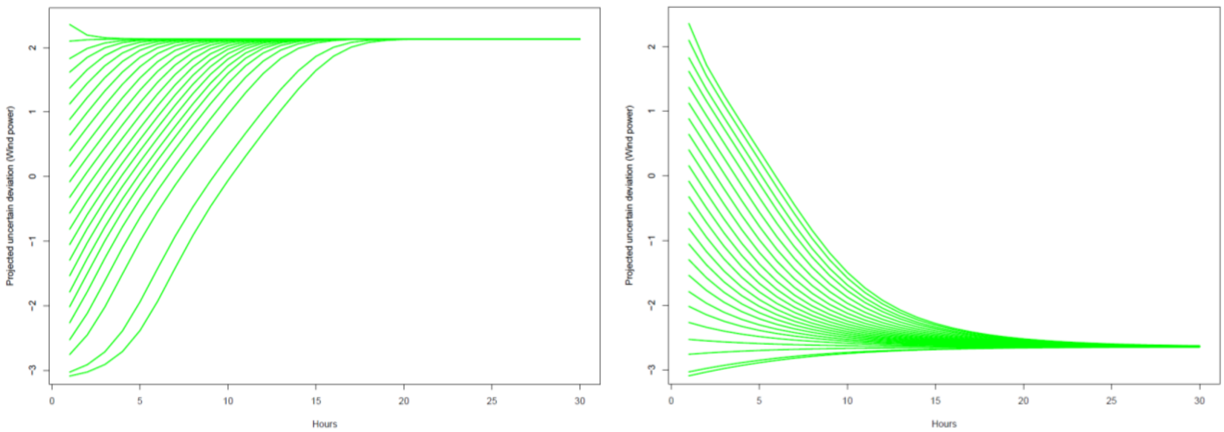

Figure 4.C.: q worst-case convergence patterns for π_{3t}^{+} and π_{3t}^{-}

The same convergence analysis can be carried out for equation **(36c)**. **Figures 5.A., 5.B.** and **5.C.** respectively plot the trajectories generated by **(36c)** corresponding to various starting values for projected demand, photovoltaic and wind capacity factors. When letting our projected parameters vary such that we maximize their absolute variations between successive time periods, we observe the convergence towards "oscillatory regimes", i.e. stable trajectories where the parameters display a cyclic pattern and the average of variations tends towards zero.

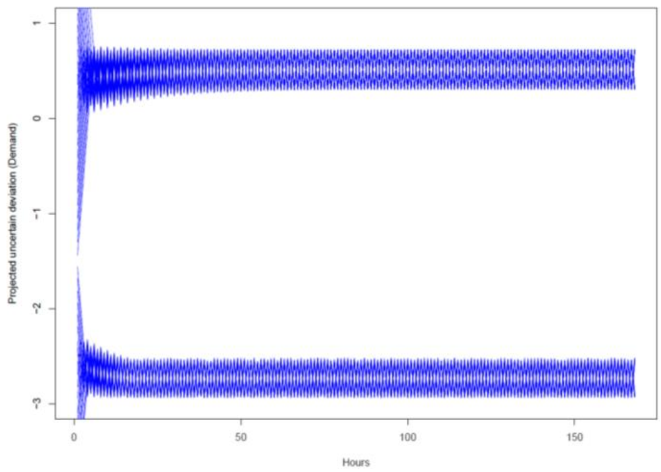

Figure 5.A.: q worst-case convergence patterns for π_{1t}^{V}

Quite strikingly, all parameters rapidly convergence towards oscillatory regimes within a small number of steps and for any starting value. Recalling that for each time step, our algorithm selects the maximum absolute variation possible conditional on the value of the uncertain parameter, there is no guarantee that the theoretical trajectory defined by an oscillatory regime actually corresponds to the theoretical trajectory maximizing variability. For instance, for π_{1t}^{V} fixed, there may exist $a \in \mathbb{R}$ such that $|a|$ is inferior to the solution $b \in \mathbb{R}$ to **(36c)** but solution to $\pi_{1t}^{V} + a$ is superior to the solution of $\pi_{1t}^{V} + b$.

The same calculations are carried for all four seasons, keeping constant the same values for parameters q, $|\mathcal{B}_i|^*$.

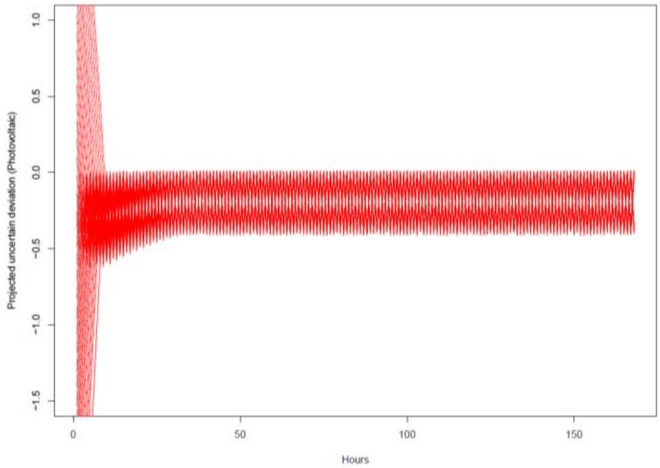

Figure 5.B.: q worst-case convergence patterns for π_{2t}^{V}

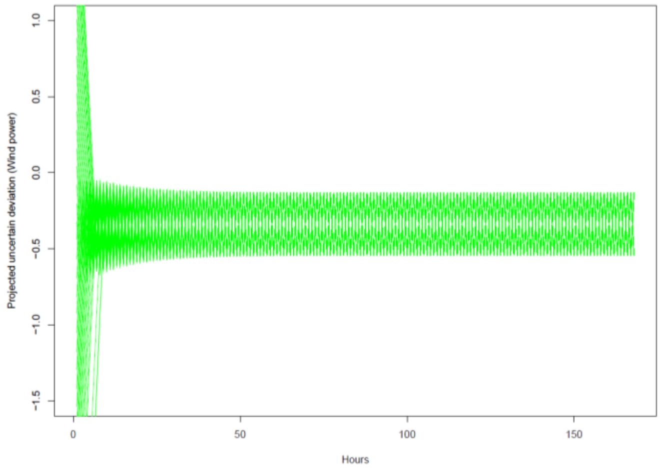

Figure 5.C.: q worst-case convergence patterns for π_{3t}^{V}

IV.B. Baseline simulation results

We compare the investment levels and cost performance of the optimal mix obtained when hedging against a variety of extreme trajectories. We respectively note **L**, **H** and **V** the trajectories for which residual demand takes its lowest, highest and most volatile values. As residual demand corresponds to the difference of electric load and renewables generation, we may reasonably assume its variability increases with renewable output. Thus, without any constraints on renewable capacity, it may be less costly to minimize investment in renewable capacity so additional costs required for enhancing system flexibility are avoided. In order to clearly observe the effects of renewable penetration on investment levels and cost performance, we constraint the wind and photovoltaic capacities to be equal to $\boldsymbol{S}$ GWe, where $\boldsymbol{S} \in \{0; 2; 4; 6\}$.

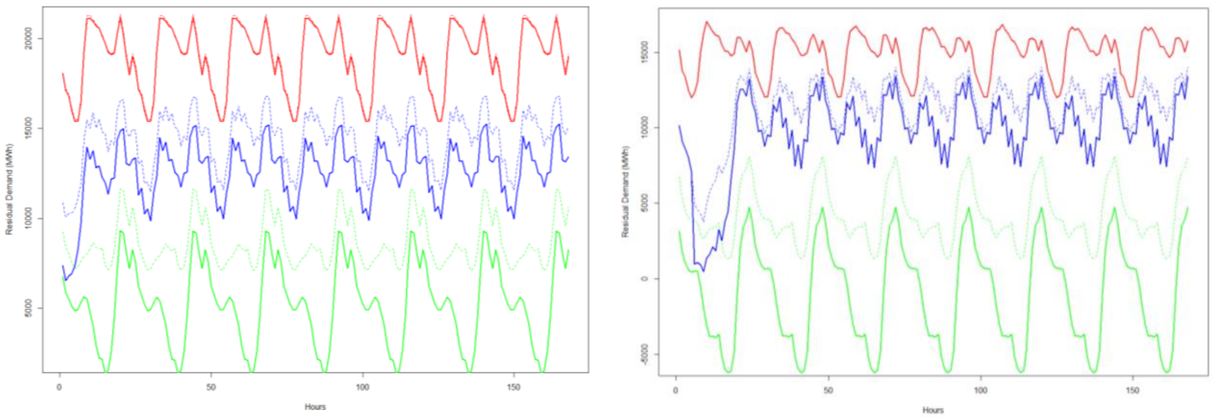

Figure 6: Extreme residual demand trajectories for Winter (left) and Spring (right), $q = 0.95$

Note: The red, green and blue plain curve corresponds to trajectory **H**, **L** and **V** for $\boldsymbol{S} = 6$. The dotted lines correspond to residual demand trajectories for $\boldsymbol{S} = 2$.

As illustrated in **Figure 6**, our method provides an interval for the set of values within which residual demand can fluctuate. In an extreme fashion, the difference between green dotted and plain lines strikingly shows how increasing renewable capacities reshapes the residual demand curve and increases the depth of the “duck dive” [13], which describes the timing imbalance between peak load in the evening and solar generation peak in the afternoon. The difference between successive peak and valley on the green plain curve translates into a ramp need of approximately 10 GWh within a few hours, for both seasons. This indicates an adequate generation mix under high renewable penetration requires production units with steep ramping capacities and low minimum production levels, associated with an active management of renewable output, including energy storage, curtailment and demand response. Finally, the rapid oscillations of the blue curves, which reflect residual demand short term variability, make visible the need for generation units with both moderate to high ramping capacities and low minimum uptime and downtime.

The optimal investment levels corresponding to a robustness level of $q = 0.95$ are reported in **Table 2.A**. First, the nuclear capacity is superior when hedging against **H** only, while the storage capacities are generally significantly lower compared to the **L+H** and **L+H+V** cases investment levels. Second, the optimal investment capacity decreases with renewable penetration, while the sum of CCGT and GT capacities increase in almost all cases. This suggests a higher renewable penetration requires higher generation flexibility from conventional plants, which translates into higher capacities of peaking units with low minimum generation level and high ramping rates.

$q = 0.95$	H	L+H	L+H+V	H	L+H	L+H+V
	$S = 0$			$S = 2$		
Combined cycle gas turbine	2.7	12.15	9.45	2.7	12.15	10.8
Gas turbine	0	0	2.7	0	0.9	0
Nuclear	19.2	9.6	9.6	19.2	9.6	12.8
Wind	0	0	0	2	2	2
PV	0	0	0	2	2	2
Battery storage	3663.06	8472.327	8454.993	3541.751	3392.070	6388.972
	$S = 4$			$S = 6$		
Combined cycle gas turbine	6.75	14.85	10.8	5.4	12.15	12.15
Gas turbine	0	0.9	1.8	0	1.8	0.9
Nuclear	16	6.4	9.6	16	9.6	9.6
Wind	4	4	4	6	6	6
PV	4	4	4	6	6	6
Battery storage	2890.556	5030.130	9575.962	7302.519	3534.021	6928.550

Table 2.A.: Optimal investment level by technology for various levels of renewables capacity (in GWe)

Note: For any level of $\boldsymbol{S}$, the investment levels corresponding to column **H** (resp. **L+H**) are obtained when hedging against the highest residual demand trajectory (resp. when hedging both against the lowest and highest residual demand trajectories).

Finally, no clear pattern regarding the complementarity or substitutability of storage and peaking units can be identified. While CCGT and storage capacities move in opposite directions in the **H** case, a weak positive correlation may be observed for the **L+H** and **L+H+V** cases for $\boldsymbol{S} \geq 2$ GWe. This complementarity is explained by the fact that, while higher capacities in battery storage allow the smoothing of highly variables renewable production when capacities increase, thermal peaking technologies remain necessary as the capacity factors of wind and solar units is quasi null in the level-maximizing RD trajectory. This pattern of substitutability and complementarity of peaking units and storage, conditional on residual demand trajectories, generates a non-linear relationship between the level of renewable penetration and the required level of peaking capacities.

$q = 0.95$	**H**				**L+H**				**L+H+V**			
$S = 0$												
Total investment costs (B€)	75.48				47.07				46.11			
Average unit cost (week **H**, €/MWh)	9.907	8.282	8.282	12.01	35.33	27.74	33.51	36.56	38.20	27.73	33.54	43.72
Average unit cost (week **L**, €/MWh)	8.359	8.361	9.916	8.303	12.14	8.322	8.373	8.303	12.15	8.322	8.373	8.303
Average unit cost (week **V**, €/MWh)	8.296	8.304	8.286	8.330	28.78	20.31	12.37	16.35	28.77	20.30	12.37	16.35
$S = 2$												
Total investment costs (B€)	79.65				50.48				61.76			
Average unit cost (week **H**, €/MWh)	9.817	8.282	7.928	12.03	35.15	27.73	30.76	36.56	26.01	16.12	21.02	27.78
Average unit cost (week **L**, €/MWh)	6.971	6.116	7.966	7.043	7.847	6.611	8.674	6.686	6.973	5.668	6.541	5.500
Average unit cost (week **V**, €/MWh)	7.784	7.810	7.775	8.130	24.98	16.97	9.642	15.25	15.42	7.757	7.775	8.094
$S = 4$												
Total investment costs (B€)	84.28				44.54				54.77			
Average unit cost (week **H**, €/MWh)	16.87	8.282	10.03	19.14	44.11	39.39	37.76	45.78	35.01	27.73	28.02	38.69
Average unit cost (week **L**, €/MWh)	5.626	8.621	19.36	6.001	9.889	6.673	15.37	4.533	5.626	4.016	7.929	4.058
Average unit cost (week **V**, €/MWh)	7.292	7.301	7.268	7.945	32.44	26.67	21.03	28.62	21.14	13.54	7.248	14.70
$S = 6$												
Total investment costs (B€)	77.82				58.67				58.88			
Average unit cost (week **H**, €/MWh)	16.57	8.282	7.606	19.10	34.79	27.73	25.27	36.56	34.79	27.73	25.27	36.56
Average unit cost (week **L**, €/MWh)	4.666	26.86	39.44	17.33	4.823	29.58	45.18	20.43	4.453	27.07	39.31	17.69
Average unit cost (week **V**, €/MWh)	6.780	6.885	6.735	7.694	17.54	12.22	7.033	12.40	17.49	10.93	6.736	11.81

Table 2.B.: Total investment costs and average unit cost by worst-case trajectory for various levels of renewables capacity (in GWe)

Note: For each column **H**, **L+H** and **L+H+V**, sub-columns respectively correspond to the average unit cost obtained for Winter, Spring, Summer and Autumn.

Table 2.B. summarizes the average unit generation cost, in €/MWh, and the total investment cost, in billion €, associated to each optimal mix presented in **Table 2.A**. When hedging against **H** trajectory only, the significantly higher nuclear capacity entails consistently higher investment costs compared to the **L+H** and **L+H+V** investment cases. Yet, as we assume nuclear plants to have a higher minimum generation level and uptime/downtime compared to peaking units, mixes with significant nuclear capacities may not be flexible enough to

accommodate large residual demand fluctuations if they are not adapted to renewable capacities. The minimum generation level may be too high to keep plants running during periods with high renewable generation, leading to generation surpluses, curtailment or even infeasibilities.

Surprisingly, in terms of average unit cost, the optimal mixes obtained when hedging against **H** only consistently outperforms the more flexible mixes obtained in the **L+H** and **L+H+V** cases and for all three types of worst-case trajectories. Although the average unit cost for trajectory **H** increases for all seasons, it remains almost 60 % lower on average. While the average unit cost for trajectory **L** are higher for low renewable penetration levels compared to the **L+H** and **L+H+V** investment cases, this difference seems to vanish for high renewable capacities.

However, we may expect that a higher share of nuclear capacities may result in a higher proportion of curtailment for trajectory **L**, which is not confirmed by our simulations. For $\boldsymbol{S} = 6$, 17.3%, 17% and 7% of wind generation are curtailed in Spring, Summer and Autumn respectively. Concerning photovoltaic output, 29.5% of production is curtailed in Spring, 38.2% in Summer and 28.3% in Autumn. By comparison, respectively 22.3%, 19.9% and 9.1% of wind output is curtailed for the **L+H** investment case, while 26.4%, 43.8% and 32.4% of photovoltaic generation is curtailed. Finally, while the **L+H+V** investment case is by construction the most flexible generation mix for any level of renewable penetration, 16.6%, 15.7% and 7.3% of wind generation is curtailed and photovoltaic curtailment levels remain above 30%. In theory, more flexible generation mixes with higher peaking capacities would result in lower curtailment levels as they can better accommodate rapid fluctuations of renewable output than nuclear plants. However, curtailment may be the less costly option overall, as renewable output may require a succession of start-ups and shut-downs to accommodate its variability, which are assumed to be very costly for CCGT and GT technologies. In conclusion, we observe a clear trade-off between high variable costs and low fixed costs mixes depending on the targeted level of system flexibility.

However, comparing generation mixes in terms of cost performance over a sample of extreme scenario, with a potentially low probability of occurrence, has little relevance in terms of average cost performance. In order to approximate the yearly distribution of production cost associated with a given generation mix, we approximate the yearly Net Load Duration Curve (NLDC) with a 4-week sample drawn from a set of 52 weeks, corresponding to a full year of demand and renewable capacity factor data. We then scale up the sample so it matches the number of hours contained in one year, and sort the resulting residual demand series by decreasing order to we obtain an Approximate NLDC. This method is introduced in [14]. We select the sample of weeks (one week for each season) that minimizes the distance between the NLDC and Approximate NLDC, measured in terms of Root-Mean-Square-Error (RMSE) and normalized RMSE. Using 2018 as a reference year for our cost analysis, our optimal 4-week sample has a RMSE of 93.53 and a normalized RMSE of 0.46 %, which corresponds to a highly accurate approximation. The approximate average yearly unit cost of generation are presented in **Table 2.C**.

$q = 0.95$	H	L+H	L+H+V
$S = 0$			
Approximate average unit cost (€/MWh)	8.320	20.41	20.83
$S = 2$			
Approximate average unit cost (€/MWh)	7.836	17.97	9.994
$S = 4$			
Approximate average unit cost (€/MWh)	7.417	27.15	14.74
$S = 6$			
Approximate average unit cost (€/MWh)	6.881	13.00	12.61

Table 2.C.: Approximate yearly average unit cost for various levels of renewables capacity (in GWe)

Again, for all levels of $\boldsymbol{S}$, the optimal mix corresponding to the **H** investment cases outperforms other mixes in terms of cost performance. Moreover, while the approximate average unit cost decreases with renewable penetration, no clear trend can be observed for the **L+H** and the **L+H+V** investment cases. The latter performs better than the **L+H** case, but remains on average twice as costly as the **H** investment cases.

Overall, optimal generation mixes obtained when hedging against all three types of worst-case trajectories systematically exhibit lower investment costs, as they include a high share of peaking units with low overnight costs. However, these mixes systematically exhibit higher average unit costs, both for extreme and representative trajectories, in addition to higher curtailment levels. In the absence of CO2 emission targets, our results entail a trade-off between these two categories of costs, which depends on the life-duration of generation units, probability of occurrence of extreme weeks and the distribution of costs between consumer categories.

IV.C. Sensitivity analysis

IV.C.1. Sensitivity to nuclear technical assumptions

As stated above, ramping constraints are not binding using an hourly time step. While extremely volatile residual demand trajectories remain feasible with a large share of nuclear power in the mix, this may not be the case anymore using a minute time step. Moreover, as underlined in [15] **Cany et al. (2018)**, increasing renewable penetration increase the frequency of extreme nuclear power ramps and annual required shut-downs/start-up events.

Even though nuclear plants can technically be operated in load-following mode [16], nuclear plants are traditionally operated in base load mode with low to moderate output variations. In order to capture the effect of this operation mode on investment and dispatching decisions, we constrain the absolute nuclear power ramp of each individual plant, denoted $\boldsymbol{r_{NUC}}$, to be inferior to 25% and 15% of the difference of the maximum and minimum generation level. The optimal investment and cost performance results are shown in **Table 3.A.**, **3.B**, **3.C.** and **Table 4.A.**, **4.B. 4.C.** respectively.

A comparison between **Table 3.A.** and **Table 2.A.** immediately shows that, for investment case **H**, the capacity of CCG and GT units increases for any level of renewable penetration. This confirms that nuclear units were previously used in load-following or peaking mode. Except for null values of renewable penetration, the installed capacities of storage and peaking units are also generally higher for the **L+H** and **L+H+V** investment cases. For a given investment case, a clear negative correlation between peaking units and storage capacities can be noted. This negative correlation is even more marked in **Table 4.A.** in Appendix, which implies a decrease in the degree of substitutability of storage and peaking units. This suggests that the direction of the elasticity of substitution between storage and peaking technologies (CCG and GT) depends of the technical characteristics of other technologies available in the mix. More specifically, this suggests the elasticity of substitution between storage and peaking units is a decreasing function of the flexibility level of alternative technologies.

$q = 0.95$ $r_{NUC} = 25\ \%$	H	L+H	L+H+V	H	L+H	L+H+V
	$S = 0$			$S = 2$		
Combined cycle gas turbine	6.75	12.150	8.1	6.75	12.150	10.8
Gas turbine	0	0	1.8	0	0.9	1.8
Nuclear	16	9.6	12.8	16	9.6	9.6
Wind	0	0	0	2	2	2
PV	0	0	0	2	2	2
Battery storage	8112.999	6219.146	7992.364	4399.080	9431.786	13850.776
	$S = 4$			$S = 6$		
Combined cycle gas turbine	6.75	12.150	14.85	6.75	12.15	12.15
Gas turbine	0	0.9	0.9	0	0.9	0.9
Nuclear	16	9.6	6.4	16	9.6	9.6
Wind	4	4	4	6	6	6
PV	4	4	4	6	6	6
Battery storage	4573.590	6419.424	6609.16	2890.556	9320.521	10627.595

Table 3.A.: Optimal investment level by technology for $r_{NUC} = 25\%$ and various levels of renewables capacity (in GWe)

This confirms the intuition that, in the absence of carbon capture technologies, the capacities of CO2 emitting technologies can only be diminished if alternative dispatchable flexible technologies are available. Demand-side management is a special case, as its elasticity of substitution with storage is a function of the correlation (and cross-correlations) of demand and renewable generation. If the correlation is negative, a share of demand during low renewable generation periods may be transferred to high generation ones with higher demand.

Moreover, the degree of substitutability of storage and peaking technologies is limited by trajectories defined by long periods of low renewable generation, during which a minimum peaking capacity remains necessary. The substitutability of storage and peaking technologies must also be evaluated by their ability to balance residual demand variations. The optimal storage capacity should indeed allow system balance for any residual demand path, which cannot be more volatile than trajectory **V** by definition. Investment in peaking units can be avoided if for any sequence, successive residual demand variations self-balance in time, i.e. if their moving average is close to zero. The order of values taken by residual demand is greatly relevant to evaluate the feasibility of any storage decision sequence. This limitation is not explicitly accounted for in our model formulation but can easily be alleviated by adding an extra dynamic constraint on the value of residual balance moving average. Finally, our methodology maximizes residual demand variability locally, conditional on residual demand value, but offers no guarantee that it maximizes residual demand variability globally. Our model allows investment decisions hedging against residual demand trajectories with maximal short-term volatility only. This limitation shall be the topic of further research.

Table 3.B. and **4.B.** (see in Appendix) both show that decreasing nuclear units ramping rate, the average unit cost is unambiguously lower for all seasons, all investment cases and all levels of renewable penetration. Decreasing nuclear flexibility increases the utilization rate of CCG and GT units, which increases the average unit cost as they have higher marginal cost. Moreover, this flexibility loss translates into higher curtailment rates for the **H** investment case compared to alternative ones. For trajectory **L** with a nuclear ramping rate of 25% and $\boldsymbol{S} = 6$, 22.8%, 13.5% and 2.2% of wind generation are curtailed in Spring, Summer and Autumn respectively. Yet, 24.2% of photovoltaic production is now curtailed in Spring, 49.2% in Summer and 46.2% in Autumn. By comparison, respectively 16.2%, 14.5% and 5.2% of wind output is curtailed for the **L+H+V** investment case, while 28%, 31.4% and 25.2% of photovoltaic generation is curtailed.

As nuclear flexibility diminishes, the costs of renewable integration become comparatively higher for a mix with a high share of nuclear capacity. In the **H** investment case, the decrease in nuclear flexibility is not compensated by hedging against low residual values trajectories, which results in higher curtailment rates for all seasons. However, for any level of renewable penetration, the economic performance of nuclear based mixes remains better, except for trajectory **L** which consistently exhibits higher average unit cost due to higher curtailment rates.

Finally, **Table 3.C.** and **Table 4.C.** in Appendix clearly show that, even when restricting nuclear units to baseload operation mode, nuclear based mixes remain on average the least expensive investment choice in terms of average unit cost.

$q = 0.95$ $r_{NUC} = 25$ %	H				L+H				L+H+V			
$S = 0$												
Total investment costs (B€)	67.25				46.69				56.82			
Average unit cost (week **H**, €/MWh)	17.09	8.282	14.10	19.06	35.33	27.73	33.51	36.56	26.20	16.13	23.77	28.69
Average unit cost (week **L**, €/MWh)	8.291	8.322	8.382	8.303	15.25	8.322	8.383	8.303	8.291	8.322	8.382	8.303
Average unit cost (week **V**, €/MWh)	8.287	8.306	8.288	8.295	28.77	20.36	12.76	16.42	17.49	8.620	8.288	8.295
$S = 2$												
Total investment costs (B€)	70.54				51.50				51.58			
Average unit cost (week **H**, €/MWh)	16.91	8.282	12.05	19.06	35.15	27.73	30.77	36.56	35.19	27.73	30.76	37.89
Average unit cost (week **L**, €/MWh)	6.972	6.258	9.511	5.502	6.949	5.185	6.637	5.502	6.949	5.180	6.412	5.502
Average unit cost (week **V**, €/MWh)	7.784	7.784	7.818	8.130	24.95	16.82	8.742	14.80	24.95	16.74	8.844	14.79
$S = 4$												
Total investment costs (B€)	74.47				54.89				44.80			
Average unit cost (week **H**, €/MWh)	16.73	8.282	9.803	19.06	34.97	27.73	28.02	36.56	44.11	39.39	37.76	45.52
Average unit cost (week **L**, €/MWh)	5.628	7.524	16.47	4.698	5.770	5.859	13.06	5.392	9.511	6.694	13.03	5.238
Average unit cost (week **V**, €/MWh)	7.292	7.303	7.312	7.914	21.14	13.60	7.623	13.27	32.44	26.63	21.08	28.62
$S = 6$												
Total investment costs (B€)	78.09				59.29				59.51			
Average unit cost (week **H**, €/MWh)	16.82	8.282	8.125	19.14	34.79	27.73	25.27	36.56	34.79	27.73	25.27	36.56
Average unit cost (week **L**, €/MWh)	5.113	30.59	46.54	21.63	8.884	25.76	33.97	15.79	4.703	25.05	36.14	38.72
Average unit cost (week **V**, €/MWh)	6.802	6.950	6.806	7.768	17.51	10.54	6.817	11.77	17.45	10.27	6.770	11.78

Table 3.B.: Total investment costs and average unit cost by worst-case trajectory for $r_{NUC} =$ 25% and various levels of renewables capacity (in GWe)

$q = 0.95$ $r_{NUC} = 25$ %	H	L+H	L+H+V
$S = 0$			
Approximate average unit cost (€/MWh)	8.443	20.93	11.51
$S = 2$			
Approximate average unit cost (€/MWh)	7.869	17.63	17.57
$S = 4$			
Approximate average unit cost (€/MWh)	7.354	14.95	27.12
$S = 6$			
Approximate average unit cost (€/MWh)	6.911	12.46	12.35

Table 3.C.: Approximate yearly average unit cost for $r_{NUC} = 25\%$ and various levels of renewables capacity (in GWe)

IV.C.2. Sensitivity to q

We can finally study the sensitivity of our results by varying the degree of robustness of our model, controlled by parameter q. By definition, for $q = 1$, the set of extreme residual demand trajectories defined by our selection method includes all theoretically possible trajectories and is thus expected to be the most conservative. The conservativeness of solutions can then controlled by decreasing q, but the proportion of infeasible trajectories will increase. The value of q can be chosen by defining a probabilistic threshold so that values of q that generate trajectories with a joint probability below this threshold are ruled out.

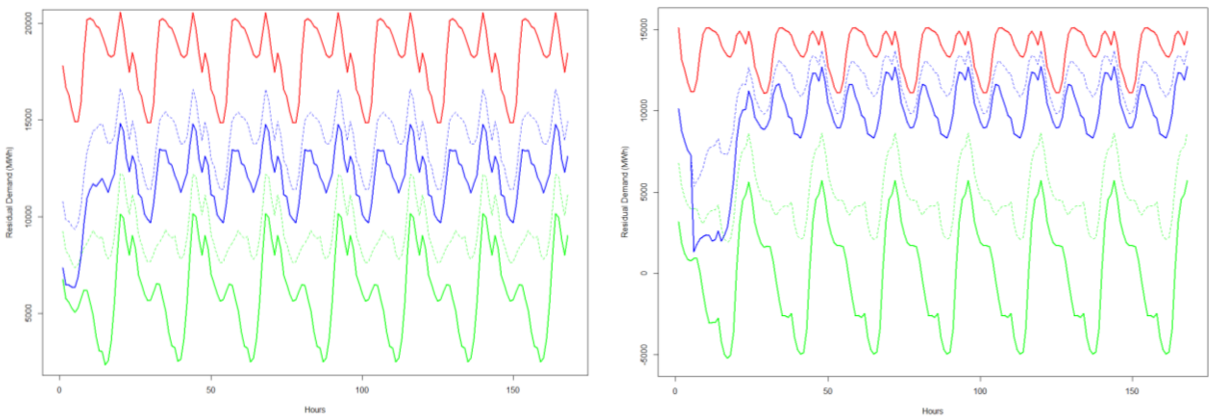

Figure 7: Extreme residual demand trajectories for Winter (left) and Spring (right), $q = 0.75$

We set $q = 0.75$ and draw in **Figure 7** the corresponding residual demand trajectories for Winter and Spring. We note little difference in terms of ramp need required for satisfying the "duck dive". Yet, for both seasons, the level of the trajectory **H** is lower while the short-term variability of the trajectory **V** is reduced. This confirms that our parameter q effectively controls both the level and variability of residual demand trajectories generated by our model.

The optimal investment levels are shown in **Table 5.A.**. We note that for the **H** investment case, the capacity of CCG turbines increases while the capacity of nuclear decreases for low renewable penetration. As the cost advantage of nuclear technology increases with the amount of generation, CCG turbines become comparatively less costly, in terms of total overnight and generation costs, as q decreases. Unsurprisingly, the total installed capacity decreases for all investment case and level of renewable penetration.

$q = 0.75$	**H**	**L+H**	**L+H+V**	**H**	**L+H**	**L+H+V**
	$S = 0$			$S = 2$		
Combined cycle gas turbine	4.050	9.45	6.75	4.050	12.15	9.45
Gas turbine	0	0	0	0	0	0.9
Nuclear	16	9.6	12.8	16	6.4	9.6
Wind	0	0	0	2	2	2
PV	0	0	0	2	2	2
Battery storage	1131.393	6680.721	10858.910	1211.035	13952.555	9866.306
	$S = 4$			$S = 6$		
Combined cycle gas turbine	4.050	13.5	9.45	4.050	10.8	9.45
Gas turbine	0	0	0.9	0	2.7	0.9
Nuclear	16	6.4	9.6	16	6.4	9.6
Wind	4	4	4	6	6	6
PV	4	4	4	6	6	6
Battery storage	4081.681	2184.314	984.991	4317.095	3287.598	10798.593

Table 5.A.: Optimal investment level by technology for $q = 0.75$ and various levels of renewables capacity (in GWe)

These observations translate into lower investment costs for all mixes, as shown in **Table 5.B.**. Simultaneously, the average unit cost associated with trajectory **L** are generally lower compared to **Table 2.B.**, while the unit cost associated with trajectory **H** are consistently higher. The lower share of nuclear and storage capacities translate into a higher utilization rate of peaking units during high demand periods, while less curtailment is necessary during periods with strong renewable generation.

This finally translates into higher approximate average unit costs for all investment cases and renewable installed capacity level. A higher degree of conservativeness, associated to a higher q, thus automatically results in higher total installed capacities and investment costs, but may result in better unit cost performance depending on the mix composition.

$q = 0.75$	H				L+H				L+H+V			
$S = 0$												
Total investment costs (B€)	64.04				44.73				55.56			
Average unit cost (week **H**, €/MWh)	14.84	7.617	14.94	12.58	32.60	22.55	33.85	28.86	23.40	11.41	24.07	20.08
Average unit cost (week **L**, €/MWh)	8.649	8.847	11.62	8.661	14.66	8.877	11.48	8.601	8.633	8.816	9.201	8.601
Average unit cost (week **V**, €/MWh)	8.462	8.379	8.837	7.783	26.95	20.64	17.15	12.86	15.68	8.372	8.837	7.765
$S = 2$												
Total investment costs (B€)	67.96				39.74				49.53			
Average unit cost (week **H**, €/MWh)	14.83	7.616	12.39	13.88	41.80	34.21	40.45	37.65	32.60	22.54	30.69	28.86
Average unit cost (week **L**, €/MWh)	7.455	6.924	26.38	6.508	22.11	5.931	6.218	5.983	7.818	5.931	10.16	5.983
Average unit cost (week **V**, €/MWh)	7.522	7.792	8.326	7.337	34.24	30.09	28.96	24.28	22.92	16.62	13.53	9.805
$S = 4$												
Total investment costs (B€)	72.35				42.68				51.94			
Average unit cost (week **H**, €/MWh)	14.51	7.598	9.554	12.20	41.66	34.21	37.31	37.57	34.22	22.54	27.57	29.45
Average unit cost (week **L**, €/MWh)	6.194	5.352	12.32	5.000	13.83	16.98	15.80	5.395	6.963	7.237	20.08	17.17
Average unit cost (week **V**, €/MWh)	6.985	7.172	7.778	6.859	30.21	26.09	25.57	21.05	19.06	13.38	11.50	9.504
$S = 6$												
Total investment costs (B€)	76.30				45.81				57.50			
Average unit cost (week **H**, €/MWh)	14.48	7.598	11.13	12.20	49.59	34.21	34.20	42.78	32.60	22.54	24.45	28.86
Average unit cost (week **L**, €/MWh)	5.548	21.31	37.95	16.00	8.731	21.93	39.72	16.96	4.970	16.60	25.43	11.21
Average unit cost (week **V**, €/MWh)	6.467	6.804	7.263	6.412	26.19	22.36	21.99	18.04	15.11	9.481	7.360	6.399

Table 5.B.: Total investment costs and average unit cost by worst-case trajectory for $q = 0.75$ and various levels of renewables capacity (in GWe)

$q = 0.75$	H	L+H	L+H+V
$S = 0$			
Approximate average unit cost (€/MWh)	8.757	20.92	11.47
$S = 2$			
Approximate average unit cost (€/MWh)	8.369	30.70	17.62
$S = 4$			
Approximate average unit cost (€/MWh)	7.360	27.22	15.64
$S = 6$			
Approximate average unit cost (€/MWh)	6.932	23.86	12.34

Table 5.C.: Approximate yearly average unit cost for $q = 0.75$ and various levels of renewables capacity (in GWe)

V. CONCLUSION

This paper thus presents an original approach to robust optimization, that does not require the definition of any uncertainty set contrary to more traditional approaches. As the dynamics of robustness have received little attention so far, we approximate the variations of uncertain parameters between successive periods using a flexible polynomial approximation method. This framework allows the estimation of a continuous approximation function, which shape and prevision interval can be tightly configured depending on the quality and distribution of the training data. This allows defining a set of limiting residual demand trajectories. However, the variability maximizing trajectory only maximizes locally the variations of residual demand. We leave for further research the definition a solution for determining the trajectory which globally maximizes residual demand trajectory over its full length.

Finally, we showed the usefulness of our enriched robust optimization method with an application to the case of Auvergne Rhône-Alpes. Hedging against trajectories with extremely high short-term variability globally increased the optimal storage and peaking capacities. While mixes with a high share of nuclear globally performed better than more adapted mixes, our results suggest that, as the stress on nuclear plants may increase with renewable penetration, the latter may become very costly and inefficient for high renewable capacities, as they may require both strong curtailment and significant peaking capacities. In the absence of any alternative non-emitting peaking technology, increasing renewable penetration while significantly decreasing CO_2 emitting technologies capacities seems unreachable.

VI. APPENDIX

Appendix to III.C :

We define the vector $\boldsymbol{\pi}^{+}_{i,0\to t} = (\pi^{+}_{i0}, \pi^{+}_{i1}, \dots, \pi^{+}_{it}) \in \mathbb{R}^{1\times(t+1)}$ and the probability threshold $\rho > 0$ such that we must have:

$$\mathbb{P}(\boldsymbol{\pi}^{+}_{i,0\to t}) = \mathbb{P}(\pi^{+}_{i0}) \prod_{t'=1}^{t} \mathbb{P}(\pi^{+}_{it'} | \pi_{it'-1}) \geq \rho \qquad \mathbf{(52)}$$

Without loss of generality, assuming that for π^{+}_{it} strictly positive, we have $\mathbb{P}(\pi^{+}_{it} | \pi_{it-1}) \leq \mathbb{P}(\pi^{+}_{it} + \epsilon | \pi_{it-1}) \Leftrightarrow \epsilon \geq 0$. We introduce an auxiliary variable $ð^{+}_{it}$ such that we have the following set of constraints for all $i \leq n$:

$$\pi^{+}_{it} = \pi^{+}_{it-1} + ð^{+}_{it} \qquad \mathbf{(53)}$$

$$\begin{aligned} ð^{+}_{it} \leq {} & \lambda_{0t} \widehat{\boldsymbol{\varphi}^{+}_{i}}\left(\boldsymbol{V}_i^{T} \pi^{+}_{it-1}; 0, \alpha, |\boldsymbol{\mathcal{B}}_i|^{*}\right) \\ & + \sum_{q \in \mathcal{Q}\setminus\{0\}} \lambda_{qt} \left(\boldsymbol{V}_i^{-T} \widehat{\boldsymbol{\varphi}^{+}_{i}}\left(\boldsymbol{V}_i^{T} \pi^{+}_{it-1}; q, \alpha, |\boldsymbol{\mathcal{B}}_i|^{*}\right) - \boldsymbol{V}_i^{-T} \widehat{\boldsymbol{\varphi}^{+}_{i}}\left(\boldsymbol{V}_i^{T} \pi^{+}_{it-1}; q - |\mathcal{Q}|^{-1}, \alpha, |\boldsymbol{\mathcal{B}}_i|^{*}\right) \right) \end{aligned} \qquad \mathbf{(54)}$$

$$\mathbb{P}(\pi^{+}_{i0}) \prod_{t'=1}^{|\mathcal{T}|} \left(1 - \sum_{q \in \mathcal{Q}\setminus\{0\}} \lambda_{qt} |\mathcal{Q}|^{-1} \right) \geq \rho \qquad \mathbf{(55)}$$

$$\lambda_{qt} \in [0; 1] \qquad \mathbf{(56)}$$

Using the monotonicity of $\mathbb{P}(\boldsymbol{\pi}^{+}_{i,0\to t})$, the probability of any given trajectory will decrease when $\boldsymbol{\pi}^{+}_{i,0\to t}$ increases. However, due to the multidimensionality of $\boldsymbol{\pi}^{+}_{i,0\to|\mathcal{T}|}$, there exists an infinity of "worst-case" vectors which saturate **(55)**. To see this, we can reformulate **(55)** as follows:

$$\mathbb{P}(\pi^{+}_{i0}) \prod_{t'=1}^{\tau} \left(1 - \sum_{q \in \mathcal{Q}\setminus\{0\}} \lambda_{qt} |\mathcal{Q}|^{-1} \right) \geq \rho \left(\prod_{t'=\tau+1}^{|\mathcal{T}|} \left(1 - \sum_{q \in \mathcal{Q}\setminus\{0\}} \lambda_{qt} |\mathcal{Q}|^{-1} \right) \right)^{-1} \qquad \mathbf{(57)}$$

An increase in the left-hand side is associated with a decrease in the right-side one. In order to define the sequence $\{ð^{+}_{it}\}_{t\in\mathcal{T}}$ that maximizes the vector length of $\boldsymbol{\pi}^{+}_{i,0\to|\mathcal{T}|}$, we solve the following linear binary optimization problem:

$$\max_{ð} \sum_{t'=1}^{|\mathcal{T}|} ð^{+}_{it}$$

s.t. **(53)-(56)** hold. Likewise, the trajectory $\{ð_{it}^{-}\}_{t\in\mathcal{T}}$ that minimizes the vector length of $\boldsymbol{\pi}_{i,0\to|\mathcal{T}|}^{-}$ is the solution to the following linear optimization problem:

$$\min_{ð} \sum_{t'=1}^{|\mathcal{T}|} ð_{it}^{-}$$

s.t. **(56)** holds and:

$$\pi_{it}^{-} = \pi_{it-1}^{-} + ð_{it}^{-} \qquad \mathbf{(58)}$$

$$ð_{it}^{-} \geq \lambda_{0t}\widehat{\boldsymbol{\varphi}_i^{-}}\left(\boldsymbol{V}_i^T\pi_{it-1}^{-}; 0, \alpha, |\boldsymbol{\mathcal{B}}_i|^*\right) + \sum_{q\in\mathcal{Q}\setminus\{0\}} \lambda_{qt}\left(\boldsymbol{V}_i^{-T}\widehat{\boldsymbol{\varphi}_i^{-}}\left(\boldsymbol{V}_i^T\pi_{it-1}^{-}; q, \alpha, |\boldsymbol{\mathcal{B}}_i|^*\right) - \boldsymbol{V}_i^{-T}\widehat{\boldsymbol{\varphi}_i^{-}}\left(\boldsymbol{V}_i^T\pi_{it-1}^{-}; q-|\mathcal{Q}|^{-1}, \alpha, |\boldsymbol{\mathcal{B}}_i|^*\right)\right) \qquad \mathbf{(59)}$$

$$\mathbb{P}(\pi_{i0}^{-})\prod_{t'=1}^{|\mathcal{T}|}\left(1 - \sum_{q\in\mathcal{Q}\setminus\{0\}} \lambda_{qt}|\mathcal{Q}|^{-1}\right) \geq \rho \qquad \mathbf{(60)}$$

In a similar fashion, we can express the trajectory $\{ð_{it}^{V}\}_{t\in\mathcal{T}}$ that maximizes the variability of the vector $\boldsymbol{\pi}_{i,0\to|\mathcal{T}|}^{V}$ such that $\mathbb{P}\left(\boldsymbol{\pi}_{i,0\to|\mathcal{T}|}^{V}\right)$ as the solution of the following quadratic optimization problem:

$$\max_{ð} \sqrt{\sum_{t'=1}^{|\mathcal{T}|} {ð_{it}^{V}}^2}$$

s.t. **(56)** holds and:

$$\pi_{it}^{V} = \pi_{it-1}^{V} + ð_{it}^{V} \qquad \mathbf{(61)}$$

$$ð_{it}^{V} \leq \lambda_{mt}\widehat{\boldsymbol{\varphi}_i^{+}}\left(\boldsymbol{V}_i^T\pi_{it-1}^{+}; m, \alpha, |\boldsymbol{\mathcal{B}}_i|^*\right) + \sum_{q\in\mathcal{Q}\setminus\mathcal{Q}_m^{-}} \lambda_{qt}\left(\boldsymbol{V}_i^{-T}\widehat{\boldsymbol{\varphi}_i^{+}}\left(\boldsymbol{V}_i^T\pi_{it-1}^{V}; q, \alpha, |\boldsymbol{\mathcal{B}}_i|^*\right) - \boldsymbol{V}_i^{-T}\widehat{\boldsymbol{\varphi}_i^{+}}\left(\boldsymbol{V}_i^T\pi_{it-1}^{V}; q-|\mathcal{Q}|^{-1}, \alpha, |\boldsymbol{\mathcal{B}}_i|^*\right)\right) \qquad \mathbf{(62a)}$$

$$ð_{it}^{V} \geq \lambda_{mt}\widehat{\boldsymbol{\varphi}_i^{-}}\left(\boldsymbol{V}_i^T\pi_{it-1}^{-}; m, \alpha, |\boldsymbol{\mathcal{B}}_i|^*\right) + \sum_{q\in\mathcal{Q}\setminus\mathcal{Q}_m^{+}} \lambda_{qt}\left(\boldsymbol{V}_i^{-T}\widehat{\boldsymbol{\varphi}_i^{-}}\left(\boldsymbol{V}_i^T\pi_{it-1}^{V}; m+q, \alpha, |\boldsymbol{\mathcal{B}}_i|^*\right) - \boldsymbol{V}_i^{-T}\widehat{\boldsymbol{\varphi}_i^{-}}\left(\boldsymbol{V}_i^T\pi_{it-1}^{V}; m+q-|\mathcal{Q}|^{-1}, \alpha, |\boldsymbol{\mathcal{B}}_i|^*\right)\right) \mathbf{(62b)}$$

$$\mathbb{P}(\pi_{i0}^{V})\prod_{t'=1}^{|\mathcal{T}|}\left(\left(m - \left(\sum_{q\in\mathcal{Q}\setminus\mathcal{Q}_m^{-}} \lambda_{qt}|\mathcal{Q}|^{-1} + \sum_{q\in\mathcal{Q}\setminus\mathcal{Q}_m^{+}} \lambda_{qt}|\mathcal{Q}|^{-1}\right)\right)\right) = \rho \qquad \mathbf{(63)}$$

$$m - \left(\sum_{q \in \mathcal{Q} \setminus \mathcal{Q}_m^-} \lambda_{qt} |\mathcal{Q}|^{-1} + \sum_{q \in \mathcal{Q} \setminus \mathcal{Q}_m^+} \lambda_{qt} |\mathcal{Q}|^{-1} \right) \geq 0 \qquad \textbf{(64)}$$

Where we define the subsets $\mathcal{Q}_m^+ \subset \mathcal{Q}$ and $\mathcal{Q}_m^- \subset \mathcal{Q}$ such that $\mathcal{Q}_m^+ \cap \mathcal{Q}_m^- = \emptyset$ and $\mathcal{Q}_m^+ \cup \mathcal{Q}_m^- \cup q_m = \mathcal{Q}$, with q_m corresponding to the median. As $\lambda_{qt} = 0$ for the subset of quantiles that does not maximize the objective function in period $t \in \mathcal{T}$, **(64)** may be redundant. Yet, in order to ensure proper behavior of the model, we further impose **(64)** so that $ð_{it}^V$ is always associated to a positive probability.

Appendix to IV:

Technology	Minimum generation level (% nominal power)	Ramping rate (% of nominal power/min)	Minimum uptime/downtime (hours)
Combined cycle gas turbine	20	20	0
Gas turbine	15	8	2
Nuclear	50	2-5	10
Hydroelectric	5	15	0.1

Table 1.A.: Flexibility characteristics of main generation technologies.

Sources: Gonzalez-Salazar et al. (2018), IAEA (2018), Schill et al. (2016), IEA (2015), Schröder et al. (2013), EC JRC (2010).

Technology	Overnight cost (€/kWe)	Annual fixed & maintenance costs (€/kWe)	Unit variable cost (€/MWh)	Unit starting cost (€/MWh)	Average lifetime (years)
Combined cycle gas turbine	754	20	45	235	30
Gas turbine	400	6.4	135	542.8	30
Nuclear	3800	137	8	90	60
Hydroelectric	3200	32	0	5	80
Photovoltaic	669	19	0	0	25
Wind power	1284	45	0	0	20
Battery storage	169	5.1	0	0	10

Table 1.B.: Cost assumptions for generation technologies for 2021.

Sources: Le Cout des ENR en France, ADEME (2016) ; CRE (2018) ; "Coûts et rentabilité du grand photovoltaïque en métropole continentale", CRE (2019) ; IEA (2015) ; "Current and Prospective Costs of Electricity Generation until 2050", DIW (2013) ; OECD/IEA-NEA (2015)

Appendix to IV.A.1 :

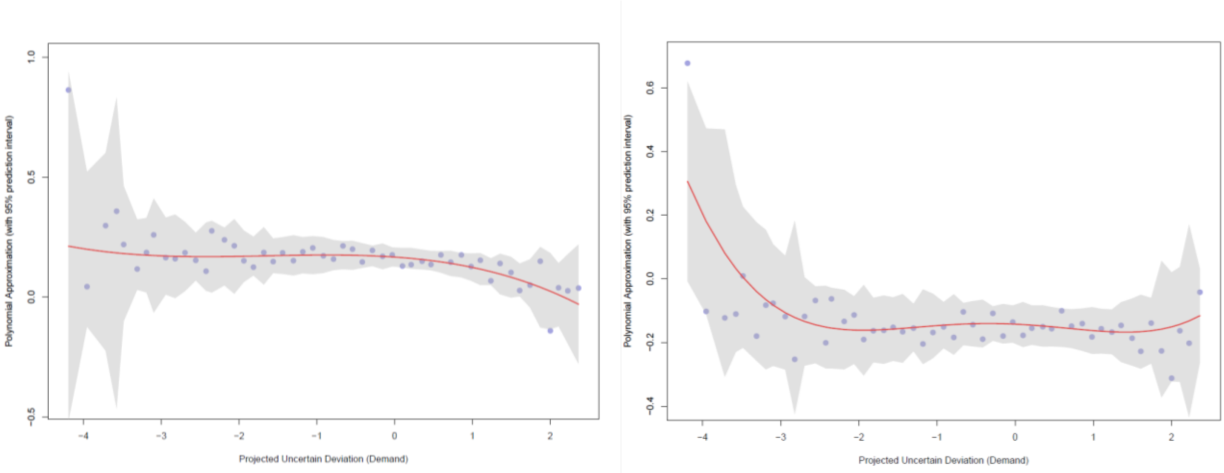

Figure 3.A.: Upper and lower quantile polynomial approximation for electricity demand for $q = 0.75$

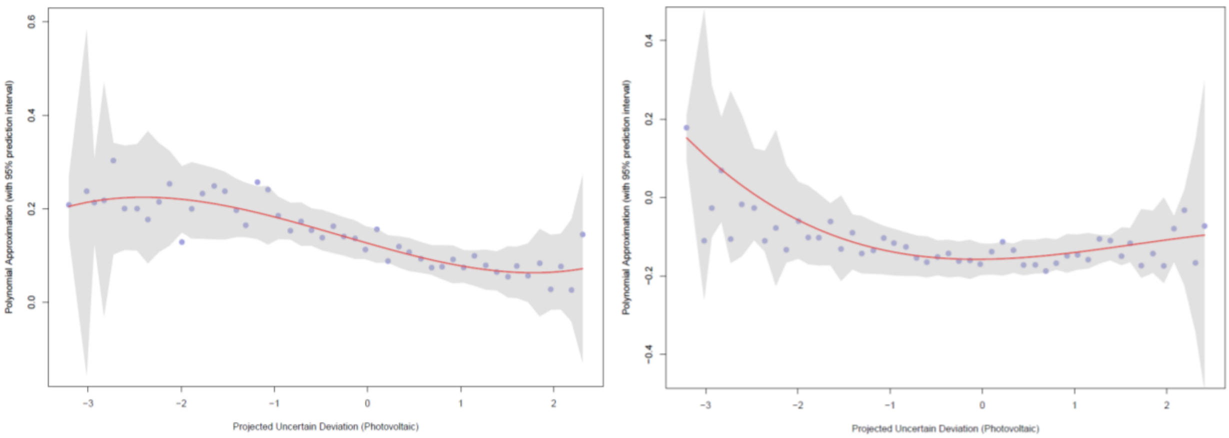

Figure 3.B.: Upper and lower quantile polynomial approximation for photovoltaic capacity factor for $q = 0.75$

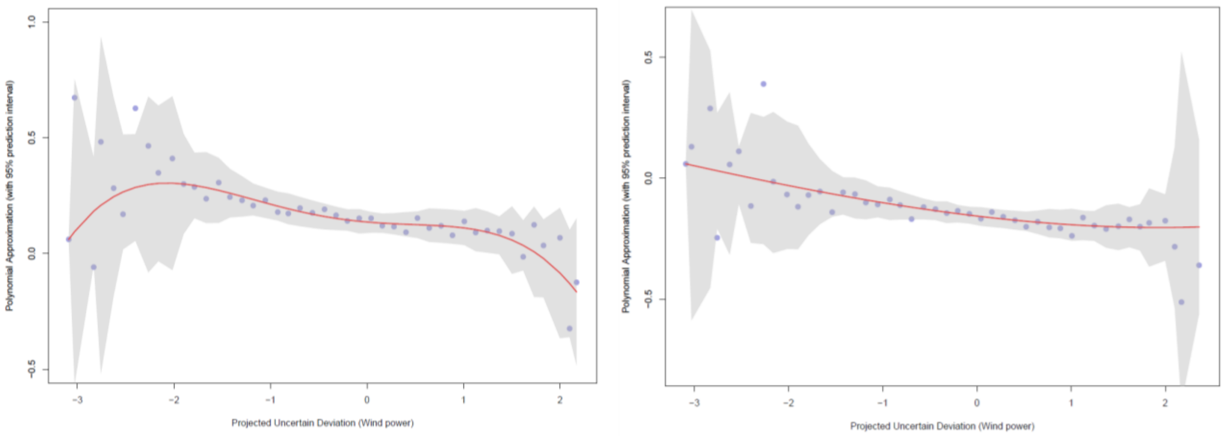

Figure 3.C.: Upper and lower quantile polynomial approximation for wind capacity factor for $q = 0.75$

Appendix to IV C.1. :

$q = 0.95$ $r_{NUC} = 15\ \%$	H	L+H	L+H+V	H	L+H	L+H+V
	$S = 0$			$S = 2$		
Combined cycle gas turbine	2.7	12.15	8.1	6.75	12.150	14.85
Gas turbine	0	0	0	0	0.9	0
Nuclear	19.2	9.6	12.8	16	9.6	6.4
Wind	0	0	0	2	2	2
PV	0	0	0	2	2	2
Battery storage	3743.083	10434.345	11884.191	2890.556	4511.516	7044.852
	$S = 4$			$S = 6$		
Combined cycle gas turbine	2.7	12.150	10.8	5.4	14.850	13.5
Gas turbine	0	0.9	0.9	0	1.8	2.7
Nuclear	19.2	9.6	9.6	16	6.4	6.4
Wind	4	4	4	6	6	6
PV	4	4	4	6	6	6
Battery storage	2444.324	5503.534	8590.854	5035.049	2277.660	2335.401

Table 4.A.: Optimal investment level by technology for $r_{NUC} = 15\%$ and various levels of renewables capacity (in GWe)

$q = 0.95$ $r_{NUC} = 15$ %	**H**				**L+H**				**L+H+V**			
$S = 0$												
Total investment costs (B€)	75.63				47.70				56.75			
Average unit cost (week **H**, €/MWh)	9.899	8.282	8.282	12.05	35.33	27.73	33.51	36.56	26.20	16.16	23.77	27.90
Average unit cost (week **L**, €/MWh)	8.320	8.373	9.799	8.306	24.93	8.389	8.400	8.306	8.291	8.324	8400	8.306
Average unit cost (week **V**, €/MWh)	8.298	8.375	8.366	11.35	28.77	20.25	12.61	16.34	17.49	8.297	8.291	8.295
$S = 2$												
Total investment costs (B€)	70.28				50.67				40.61			
Average unit cost (week **H**, €/MWh)	17.03	8.282	12.27	19.14	35.15	27.73	30.76	36.56	44.29	39.39	40.51	45.35
Average unit cost (week **L**, €/MWh)	6.975	7.020	10.42	6.557	7.524	6.079	13.05	6.178	18.63	6.340	6.943	5.510
Average unit cost (week **V**, €/MWh)	7.823	8.014	9.172	8.170	24.95	16.95	9.971	15.21	36.27	30.19	24.94	30.19
$S = 4$												
Total investment costs (B€)	83.22				54.74				54.60			
Average unit cost (week **H**, €/MWh)	9.823	8.306	7.579	12.16	34.97	27.73	28.02	36.56	35.01	27.73	28.02	38.87
Average unit cost (week **L**, €/MWh)	5.923	9.459	21.19	6.819	5.635	11.44	15.66	4.462	5.883	5.240	9.732	4.499
Average unit cost (week **V**, €/MWh)	7.303	7.398	7.354	7.973	21.14	13.68	7.955	13.44	21.29	13.70	7.525	13.27
$S = 6$												
Total investment costs (B€)	77.44				48.34				47.69			
Average unit cost (week **H**, €/MWh)	16.56	8.282	7.900	19.10	43.93	39.39	35.01	45.69	44.81	39.39	35.01	50.60
Average unit cost (week **L**, €/MWh)	5.346	28.98	60.54	19.42	7.136	28.74	41.88	19.00	9.053	46.11	77.82	22.24
Average unit cost (week **V**, €/MWh)	6.808	6.895	6.845	7.702	28.62	23.49	20.61	27.06	28.62	25.03	17.69	27.10

Table 4.B.: Total investment costs and average unit cost by worst-case trajectory for $r_{NUC} =$ 15% and various levels of renewables capacity (in GWe)

$q = 0.95$ $r_{NUC} = 15$ %	**H**	**L+H**	**L+H+V**
$S = 0$			
Approximate average unit cost (€/MWh)	8.323	21.67	11.46
$S = 2$			
Approximate average unit cost (€/MWh)	8.324	17.87	30.70
$S = 4$			
Approximate average unit cost (€/MWh)	7.439	14.97	15.45
$S = 6$			
Approximate average unit cost (€/MWh)	6.974	23.80	23.97

Table 3.C.: Approximate yearly average unit cost for $r_{NUC} = 15\%$ and various levels of renewables capacity (in GWe)

VII. BIBLIOGRAPHY

[1] ENGELAND Kolbjørn et al. “Space-time variability of climate variables and intermittent renewable electricity production- A review”, *Renewable and Substainable Energy Reviews*, vol. 79, 2017, pp. 600-617

[2] NOURAI Ali & SCHAFER Chris, “Changing the electricity game”, *IEEE Power and Energy Magazine*, vol.7, n°4, 2009

[3] ZUGNO Marco, CONEJO Antonio J., “A Robust Optimization Approach to Energy and Reserve Dispatch in Electricity Markets”, Technical University of Denmark. Technical Report-2013, n°5

[4] SOYSTER A.L., “Convex programming with set-inclusive constraints and applications tot inexact linear programming”, *Operations Research*, vol. 21, 1973, pp. 1154-1157

[5] BABONNEAU F., VIAL J.-P., APPARIGLIATO R., “Robust Optimization for Environmental and Energy Planning” in *Uncertainty and Environmental Decision Making: A Handbook of Research and Best Practice*, 2009, pp.79-126

[6] BEN-TAL A. & NEMIROVSKI A., “Robust solutions of linear programming problems contaminated with uncertain data”, *Mathematical Programming*, vol.88, 2000, pp.411-424

[7] BERTSIMAS D. & SIM M., “Price of robustness”, *Operations Research*, vol.52, 2004, pp.35-53

[8] YUAN Yuan, LI Zukui, HUANG Biao, “Robust optimization under correlated uncertainty: Formulations and computational study”, *Computers and Chemical Engineering*, vol.85, 2016, pp.58-71

[9] JALILVAND-NEJAD Amir, SHAFAEI Rasoul, SHAHRIARI Hamid, “Robust optimization under correlated polyhedral uncertainty set”, *Computers & Industrial Engineering*, vol.92, 2016, pp.82-94

[10] LORCA Alvaro & SUN Xu Andy, “Adaptive Robust Optimization with Dynamic Uncertainty Sets for Multi-Period Economic Dispatch under Significant Wind”, *IEEE Transactions on Power Systems*, vol.30, n°4, 2015, pp.1702-1713

[11] CHEN Xin, SIM Melvyn, SUN Peng, "A Robust Optimization Perspective on Stochastic Programming", *Operations Research*, vol. 55, n°6, 2007, pp. 1058-1071

[12] NING Chao & YOU Fengqui, "Data-driven decision under uncertainty integrating robust optimization with principal component analysis and kernel smoothing methods", *Computers and Chemical Engineering*, vol. 112, 2018, pp. 190-210

[13] HOU Qingchun, ZHANG Ning, DU Ershun, MIAO Miao, PENG Fei, KANG Chongqing, "Probabilistic duck curve in high PV penetration power system: Concept, modeling, and empirical analysis in China", *Applied Energy*, vol. 242, 2019, pp. 205-215

[14] DE SISTERNES Fernando, "Risk Implications of the Deployment of Renewables for Investments in Electricity Generation", MIT, May 2014

[15] CANY C., MANSILLA C., MATHONNIERE G., DA COSTA P., "Nuclear contribution to the penetration of variable renewable energy sources in a French decarbonized mix", *Energy*, vol. 150, 2018, pp. 544-555

[16] MORILHAT Patrick, FEUTRY Stéphane, LE MAITRE Christelle, MELAINE FAVENNEC Jean, "Nuclear Power Plant flexibility at EDF", 2019, hal-01977209